Université de Montréal

Investigation of the structure and dynamics of the centromeric epigenetic mark

Par

Abbas Padeganeh

Programmes de biologie moléculaire, Faculté de médecine
Université de Montréal

Thèse présentée à la Faculté des études supérieures
en vue de l'obtention du grade de Ph.D.
en Biologie Moléculaire-Option Biologie des Systèmes

April 2014

Résumé

Le centromère est le site chromosomal où le kinetochore se forme, afin d'assurer une ségrégation fidèles des chromosomes et ainsi maintenir la ploïdie appropriée lors de la mitose. L'identité du centromere est héritée par un mécanisme épigénétique impliquant une variante de l'histone H3 nommée ***cen***tromere ***p***rotein-***A*** (CENP-A), qui remplace l'histone H3 au niveau de la chromatine du centromère. Des erreurs de propagation de la chromatine du centromère peuvent mener à des problèmes de ségrégation des chromosomes, pouvant entraîner l'aneuploïdie, un phénomène fréquemment observé dans le cancer. De plus, une expression non-régulée de CENP-A a aussi été rapportée dans différentes tumeurs humaines. Ainsi, plusieurs études ont cherchées à élucider la structure et le rôle de la chromatine contenant CENP-A dans des cellules en prolifération. Toutefois, la nature moléculaire de CENP-A en tant que marqueur épigénétique ainsi que ces dynamiques à l'extérieur du cycle cellulaire demeurent des sujets débat.

Dans cette thèse, une nouvelle méthode de comptage de molécules uniques à l'aide de la microscopie à réflexion totale interne de la fluorescence (TIRF) sera décrite, puis exploitée afin d'élucider la composition moléculaire des nucléosomes contenant CENP-A, extraits de cellules en prolifération. Nous démontrons que les nucléosomes contenant CENP-A marquent les centromères humains de façon épigénétique à travers le cycle cellulaire. De plus, nos données démontrent que la forme prénucléosomale de CENP-A, en association avec la protéine chaperon HJURP existe sous forme de monomère et de dimère, ce qui reflète une étape intermédiaire de l'assemblage de nucléosomes contenant CENP-A.

Ensuite, des analyses quantitatives de centromères lors de différenciation myogénique, et dans différents tissus adultes révèlent des changements globaux qui maintiennent la marque

épigénétique dans une forme inactive suite à la différentiation terminale. Ces changements incluent une réduction du nombre de points focaux de CENP-A, un réarrangement des points dans le noyau, ainsi qu'une réduction importante de la quantité de CENP-A. De plus, nous démontrons que lorsqu'une dédifférenciation cellulaire est induite puis le cycle cellulaire ré-entamé, le phénotype "différencié" décrit ci-haut est récupéré, et les centromères reprennent leur phénotype "prolifératif".

En somme, cet oeuvre décrit la composition structurale sous-jacente à l'identité épigénétique des centromères de cellules humaines lors du cycle cellulaire, et met en lumière le rôle de CENP-A à l'extérieur du cycle cellulaire.

Mots-clés:

Centromère protéine A, histones, nucléosomes, Chromosome, centromère, cycle cellulaire, mitose, kinetochore, différenciation des cellules souches, l'imagerie molécules uniques, épigénétique, cancer, aneuploïdie

Abstract

The centromere is a unique chromosomal locus where the kinetochore is formed to mediate faithful chromosome partitioning, thus maintaining ploidy during cell division. Centromere identity is inherited via an epigenetic mechanism involving a histone H3 variant, called ***cen***tromere ***p***rotein-***A*** (CENP-A) which replaces histone H3 in centromeric chromatin. Defects in the centromeric chromatin can lead to missegregation of chromosomes resulting in aneuploidy, a frequently observed phenomenon in cancer. Moreover, deregulated CENP-A expression has also been documented in a number of human malignancies. Therefore, much effort has been devoted to uncover the structure and role of CENP-A-containing chromatin in proliferating cells. However, the molecular nature of this epigenetic mark and its potential dynamics during and outside the cell cycle remains controversial.

In this thesis, the development of a novel single-molecule imaging approach based on total internal reflection fluorescence and the use of this assay to gain quantitative information about the molecular composition of CENP-A-containing nucleosomes extracted from proliferating cells throughout the cell cycle as well as the dynamics and cellular fate of CENP-A chromatin in terminal differentiation are described.

Here, we show that octameric CENP-A nucleosomes containing core Histones H2B and H4 epigenetically mark human centromeres throughout the cell cycle. Moreover, our data demonstrate that the prenucleosomal form of CENP-A bound by the chaperone HJURP transits between monomeric and dimeric forms likely reflecting intermediate steps in CENP-A nucleosomal assembly.

Moreover, quantitative analyses of centromeres in myogenic differentiation and adult mouse tissue sections revealed that centromeres undergo global changes in order to retain a

minimal CENP-A epigenetic code in an inactive state, upon induction of terminal differentiation. These include a robust decrease in the number of centromeric foci, subnuclear rearrangement as well as extensive loss of CENP-A protein. Interestingly, we show that forced dedifferentiation under cell cycle reentry permissive conditions, rescued the above-mentioned phenotype concomitantly with the restoration of cell division.

Altogether, this work delineates the structural basis for the epigenetic specification of human centromeres during the cell cycle and sheds light on the cellular fate of the CENP-A epigenetic code outside the cell cycle.

Key words:

Centromere protein-A, Histones, Nucleosomes, Chromosome, Centromere, cell cycle, mitosis, Kinetochore, stem cell differentiation, single-molecule imaging, epigenetics, cancer, aneuploidy

Table of Contents

Table of figures:

Chapter 1:

Chapter 2:

Chapter 3:

Chapter 4:

Abbreviations:

CENP-A: Centromere protein A

Cse4: Chromosome segregation 4

CID: Centromere identifier

CATD: CENP-A targeting domain

HCP3: HoloCentric chromosome binding Protein

HFD: Histone fold domain

HJURP: Holliday Junction recognition protein

Scm3: Suppressor of chromosome missegregation 3

CBD: Cse4 binding domain

AFM: Atomic force microscopy

EM: Electrom microscopy

FCS: Fluorescence correlation spectroscopy

ChIP: Chromatin immunoprecipitation

PTM: Posttranslational modification

SiM-Pull: Single molecule pull down

PA-CNC: Photo-bleaching-assisted copy number counting

YFP: Yellow fluorescent protein

TIRF: Total internal reflection fluorescent

A.U. : arbitrary unit

eGFP: enhanced green fluorescent protein

QUBE: QUantitative Bleaching Estimation

STORM: STochastic Optical Reconstruction Microscopy

GST: Glutathione S-transferase

PBS: Phosphate buffered saline

EDTA: Ethylenediaminetetraacetic acid

IHC: Immunohistochemistry

HP1: Heterochromatin protein 1

HECCs: HEterochromatin-localized Centromere Clusters

BrdU: Bromodeoxyuridine

IF: Immunofluorescence staining

DCP: Differentiated centromere-phenotype

DM: Differentiation media

GM: Growth media

HER2: Human epidermal growth factor receptor 2

BRET: Bioluminescence resonance energy transfer

FRET: Förster/Fluorescence resonance energy transfer

hESCs: Human Embryonic Stem Cells

hIPS: Human Induced Pluripotent Stem cells

hPSCs: Human pluripotent stem cells

Dedicated to my beautiful wife "Saeideh"

and awesome "Mom & Dad"

Acknowledgements:

First and most importantly, I would like to thank my PhD supervisor, Dr Paul S Maddox who has been not only a great mentor in science and the field of cell biology, but also a person from whom I have learnt a lot of other lessons in life. Under his mentorship, I have leant how to think and view things critically, ask questions and be open to other ideas and solutions to tackle biological problems. Thank you Paul for all your support and the precious opportunity to work with your team! I have always said that joining your lab was one of the best decisions that I've ever made and I am very proud of being a part of your team!

I would also thank all past and present members of the MDX labs including our lab manager, Joel Ryan, who has always been patient and supportive with me at different stages of my projects, Amy S Maddox, Benjamin Lacroix, Jonas F Dorn, Rajesh Ranjan, Valérie De Rop, Carlos AP Descovich, Jacques Boisvert, Anne-Marie Ladouceur, Karine Bourdages, Li Zhang, and Vincent Boudreau for all the useful discussions and their support during my graduate work. I'd also like to thank Raphael Cautain whose memories shall always stay in our hearts… Undoubtedly, my doctoral committee members Dr Julie Lessard, Dr Michel Cayouette, Dr Laura Nilson and Dr Alisa Piekny played an important role in directing my project toward the right path for which I am very grateful to them.

I also would like to thank Mohamed El Ezzy, who, no matter what, supported me and with whom I have had lots of fruitful scientific discussions particularly before my pre-doctoral exam!

I would like to thank the IRIC and University of Montreal for providing outstanding facilities and financial assistance, in particular special thanks to Dr Marc Therrien, Dr Jean-Claude

Labbé, Dr Alain Verreault and Miss Vivianne Jodoin from the University of Montreal who supported me throughout my studies.

I'd like to also take this opportinity to thank my super energetic lovely wife, Saeideh, who has always stood beside me in life and in the lab keeping me motivated by her support. Her presence made every moment sweeter and allowed me to stay on the right track. I would like to thank her dear family for being there for Saeideh and me and supporting us.

And finally, I can never forget how much support my family provided me with and that it is because of them that I am who I am today: My father and mother, my brothers Ali, Milad, Javad and my sister Elham have always been a source of motivation for me since the beginning.

Chapter 1

Nucleosomal composition at the centromere: a numbers game

Abbas Padeganeh[1], Valérie De Rop[1] and Paul S. Maddox[1, 2]

[1]Institute for Research in Immunology and Cancer (IRIC), University of Montreal, H3T1J4, Montreal, Quebec, Canada

[2]Department of Pathology and Cell Biology, Faculty of Medicine, University of Montreal, H3T1J4, Montreal, Quebec, Canada

Corresponding author: Paul S. Maddox;

Running title: CENP-A nucleosomal structure

In this chapter, I present a comprehensive overview of the centromere chromatin field, which is the subject of this thesis and serves as the introduction to chapters 2, 3 and 4.

The material presented herein was published in ***Chromosome Research***. 2013 Mar;21(1):27-36.

Author contribution:

AP and PSM developed the idea of the review article. AP wrote the manuscript and prepared the images. PSM edited the manuscript.

Summary:

The Centromere is a unique chromosomal locus where the kinetochore is formed to mediate faithful chromosome partitioning, thus maintaining ploidy during cell division. Centromere identity is inherited via an epigenetic mechanism involving a histone H3 variant, called ***cen***tromere ***p***rotein-***A*** (CENP-A) which replaces H3 in centromeric chromatin. In spite of extensive efforts in field of centromere biology during the past decade, controversy persists over the structural nature of the CENP-A-containing epigenetic mark, both at nucleosomal and chromatin levels. Here, we review recent findings and hypotheses regarding the structure of CENP-A-containing complexes.

Centromeres: epigenetics and kinetochore assembly:

Centromere regions are unique in that they direct kinetochore assembly where the spindle microtubules attach and mediate chromosome segregation, thus maintaining ploidy (Cleveland et al., 2003). Defects in the centromeric chromatin may lead to missegregation of chromosomes resulting in aneuploidy, a frequently observed phenomenon in cancer (Tomonaga et al., 2003). Centromeres are highly divergent throughout evolution and even from chromosome to chromosome within a given species (Fukagawa, 2004). However, a singular conserved feature of all centromeres is the presence of a centromere-specific histone H3 variant known as ***cen***tromere ***p***rotein-***A*** (CENP-A) in centromeric nucleosomes (Figure 1) (Palmer et al., 1991; Vagnarelli et al., 2008). In the mammalian system, CENP-A-containing nucleosomes are interspersed with canonical H3 nucleosome and the whole region is flanked by pericentric heterochromatin (Allshire and Karpen, 2008). During mitosis, CENP-A nucleosomes are suggested to undergo configuration changes and form the outer layer of the centromeric region providing a platform for the formation of the mitotic kinetochore (Figure 2) (Allshire and Karpen, 2008).

CENP-A protein:

CENP-A in humans is a 140 amino acid protein with an N-terminus divergent in sequence from that of the canonical histone H3 (Yoda et al., 2000), a C-terminal ***h***istone ***f***old ***d***omain (HFD) (Sullivan et al., 1994), and a C-terminal tail required for the recruitment of kinetochore proteins such as CENP-C and –N (Guse et al., 2011). Thus, CENP-A is often regarded to as the first player in kinetochore assembly and centromere identity. Sequence

alignment shows that the HFD-containing C-terminus of CENP-A is 62% identical to that of canonical H3 (Sullivan et al., 1994).

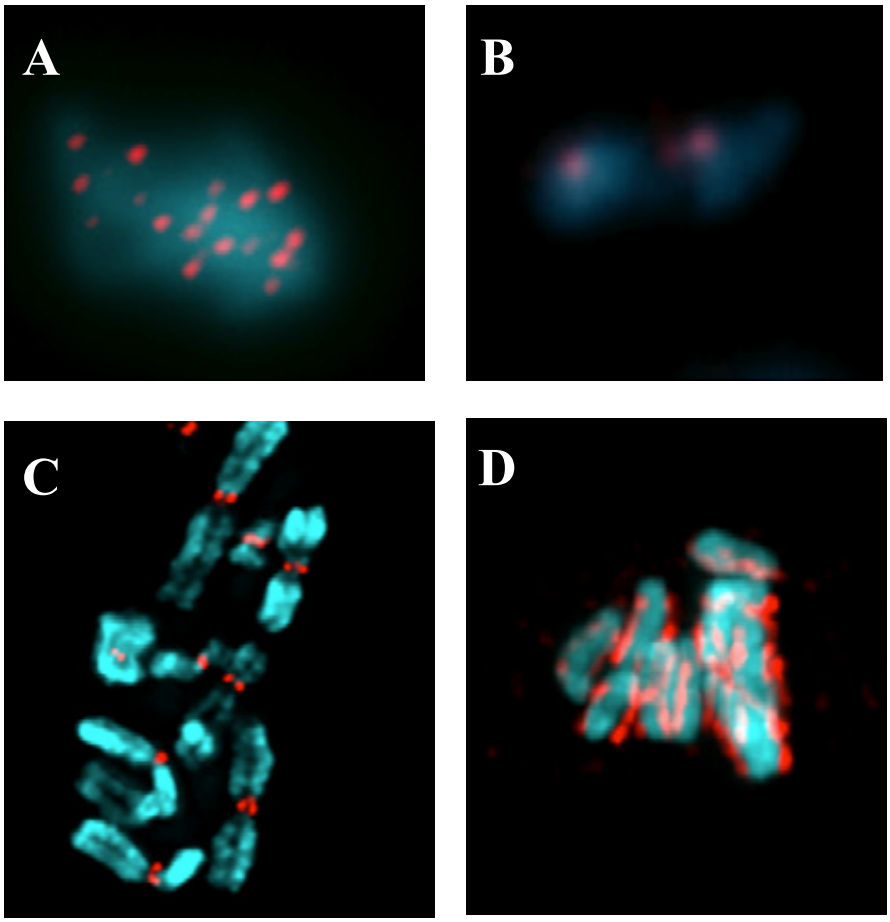

Figure 1: Presence of CENP-A is a conserved feature of centromeres in different organisms.

A. Point centromeres in *S. cerevisiae.* **B.** Chromosomes and centromeres of *S. pombe*. **C.** Regional centromeres in *H. sapiens*. **D.** Holocentric chromosmes of *C. elegans*. Blue indicates

DAPI staining of DNA. Red indicates CENP-A staining pattern on individual chromosomes (Vagnarelli et al., 2008).

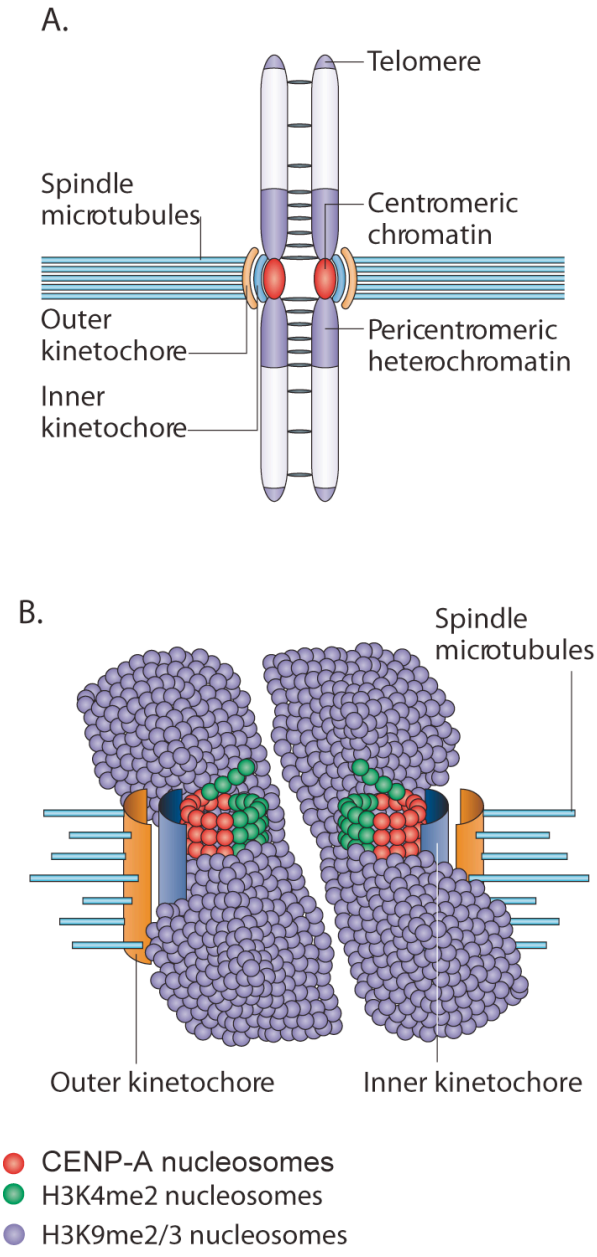

Figure 2: CENP-A-containing nucleosomes generate the chromosomal foundation upon which inner and outer kinetochore proteins are assembled.

A. The spatial relationship between CENP-A nucleosomes, pericentric heterochromatin and the mitotic kinetochore in a condensed chromosome. **B.** Close-up model for the 3D configuration of CENP-A nucleosomes relative to H3 nucleosomes and pericentric heterochromatin (Allshire and Karpen, 2008).

Within the histone fold domain, deuterium exchange measured by mass spectrometry identified a unique structural element, termed the ***C***ENP-***A*** ***t***argeting ***d***omain (CATD) comprising of the L1 and α2 helices (Black et al., 2007a). Chimeric molecules where the CATD of CENP-A is exchanged with the corresponding region in H3 (and vice versa) revealed that the CATD is a major determinant for targeting of CENP-A to the centromeres (Black et al., 2007b). Interestingly, structural data indicate that CATD also serves as an interface between CENP-A and H4 in the sub-nucleosomal $(CENP\text{-}A:H4)_2$ complex (Sekulic et al., 2010) as well as the CENP-A:CENP-A dimer in the putative octameric CENP-A nucleosomes (Bassett et al., 2012; Sekulic et al., 2010). Importantly, the CATD has been shown to interact with the ***h***olliday ***j***unction ***r***ecognition ***p***rotein (HJURP) which is an essential chaperone for CENP-A centromeric deposition (Black et al., 2004; Foltz et al., 2009; Hu et al., 2011). Despite detailed structural data discussed above, questions remain as to the precise nature of CENP-A chromatin. Complications in this regard mainly arise from inconsistencies between data obtained from biochemical work using *in vitro* nucleosome reconstitution approaches, proposing a canonical nucleosome-like octameric entity for CENP-A nucleosomes, and some lines of evidence obtained from studying endogenously purified centromeric chromatin revealing that CENP-A may not be present in a nucleosome form (octamers) and might exist in cells as some sort of tetrameric half-nucleosomes (hemisomes) at the centromere. Here we review our current understanding of the structure of CENP-A containing complexes highlighting the biological outcomes of the octamer vs. tetramer debate.

Soluble CENP-A:

Upon production of CENP-A in G2 (Howman et al., 2000; Shelby et al., 2000), the newly synthesized protein is thought to form a dimer with histone H4 and further in the cell

cycle recognized by HJURP (known as ***s***uppressor of ***c***hromosome ***m***issegregation 3, Scm3 in yeast) resulting in an equimolar complex of CENP-A:H4:HJURP (figure 3) (Cho and Harrison, 2011; Hu et al., 2011). In yeast, Scm3 binds the CATD of Cse4 (CENP-A homolog), and α2 and α3 of H4 via its ***C***se4 ***b***inding ***d***omain (CBD), with key residues conserved in HJURP (Zhou et al., 2011).

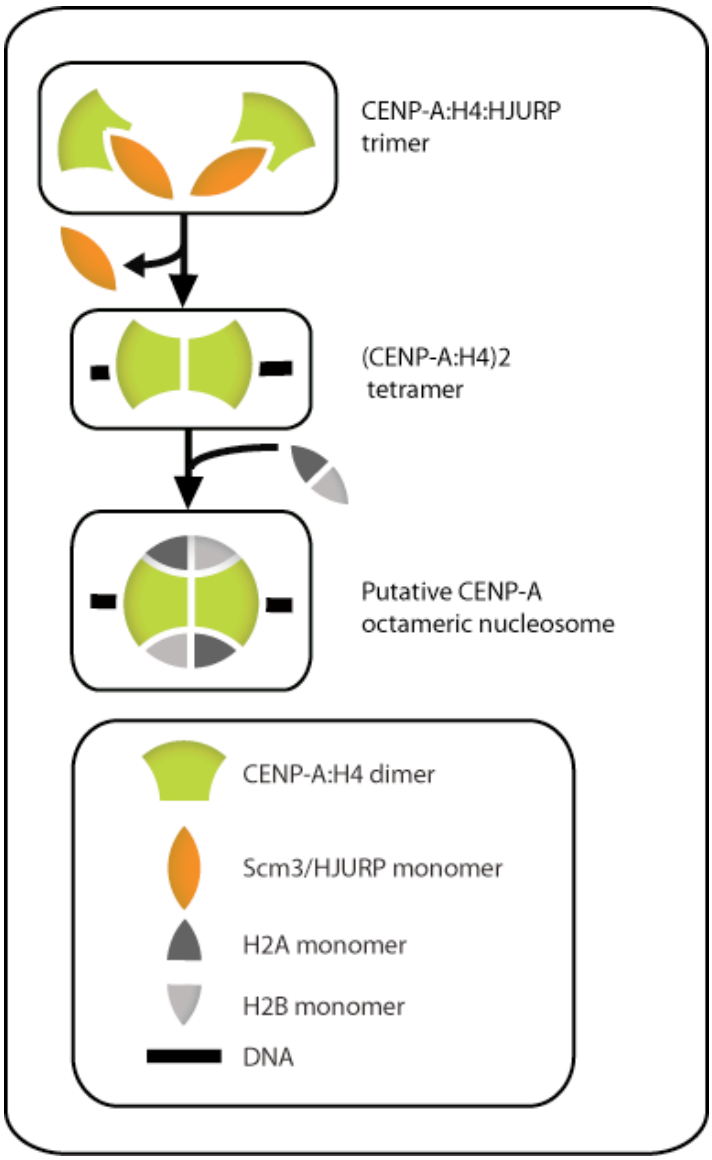

Figure 3: Proposed mechanism for the formation of putative CENP-A octameric nucleosomes.

Proposed mechanism for the formation of putative CENP-A octameric nucleosomes. Newly synthesized CENP-A along with histone H4 is suggested to bind HJURP to form a pre-nucleosomal trimeric complex. Next, in order to interact with DNA, HJURP has to be released leaving a (CENP-A:H4)$_2$ tetramer. Addition of H2A:H2B dimers then complete the octamer formation.

This interaction of HJURP with CENP-A is required to stabilize CENP-A as depletion of HJURP in human cells results in dramatically decreased CENP-A protein levels (Dunleavy et al., 2009; Foltz et al., 2009; Shuaib et al., 2010). Structural data suggest that this complex (Cse4:H4:Scm3) is not capable of interacting with DNA due to the induction of major conformational alterations in Cse4 and H4 (e.g. displacement of DNA-binding Loop 2 of H4) (Zhou et al., 2011). Moreover, it has been suggested that the presence of Scm3 in the pre-nucleosomal complex prevents the sub-nucleosomal $(Cse4:H4)_2$ tetramer formation (figure 3), a step required for the nucleosome assembly (Zhou et al., 2011). Therefore, it is intuitive to assume that Scm3 needs be recognized by another/other component(s) in order to bind the chromatin and that it has to be removed for stable incorporation of CENP-A on to centromeres.

CENP-A nucleosomal structure

The octamer model:

The crystal structure of the human CENP-A-containing nucleosome reconstituted *in vitro* from bacterially purified histones indicates homotypic octamers containing two copies of each histone molecule (Tachiwana et al., 2011). This study also revealed key features of CENP-A nucleosomes distinguishing them from canonical H3 nucleosomes. For instance, CENP-A contains a shorter αN helix lacking a key Arginine in position 49, which is an essential amino acid for DNA interaction. These findings are consistent with the data obtained independently from stepwise assembly of CENP-A nucleosomes not only confirming the octameric structure of CENP-A nucleosomes but also the loosening of the interaction between DNA superhelical termini and CENP-A (Conde e Silva et al., 2007; Panchenko et al., 2011).

CENP-A octamers formed *in vitro* have also been reported to induce conventional left-handed negative supercoiling to DNA (Barnhart et al., 2011; Conde e Silva et al., 2007; Panchenko et al., 2011; Tachiwana et al., 2011; Yoda et al., 2000). It was recently demonstrated that the mutation of the putative CENP-A:CENP-A dimer interface can abrogate centromeric targeting of CENP-A in *Drosophila* and mammalian tissue culture cells (Bassett et al., 2012; Zhang et al., 2012). In agreement with an octamer, over-expression of Cse4 (the CENP-A homolog) in budding yeast was reported to result in misincorporation of octamer-sized nucleosomes in chromosome arms (Camahort et al., 2009). These observations, along with the crystal structures available, provide solid evidence supporting the existence of octameric CENP-A nucleosomes at the centromere (Figure 6).

The Tetramer (hemisome) model:

In an effort to determine the native *in vivo* form of CENP-A chromatin, various purification and analysis techniques have been employed. One of the most extensive efforts has focused on nucleosome cross-linking followed by imunoprecipitation and ***a***tomic ***f***orce ***m***icroscopy (AFM) to investigate CID-containing nucleosomes (CID for ***c***entromere ***id***entifier, a *Drosophila* homologue of CENP-A). Challenging the octameric nucleosome concept, AFM data revealed that the height of the CID-containing interphase chromatin is half the height of canonical H3 nucleosomes (approximately 1 nm vs 2 nm) (Dalal et al., 2007). Moreover, in the beads-on-a-string structure of CID chromatin, the linker DNA is reported to be 2-3 times longer than that of conventional nucleosomes (Dalal et al., 2007). Surprisingly, the electrophoretic behavior of the purified CID-nucleosomal core particles corresponds to the presence of only one copy of each histone. This composition (CID:H4:H2A:H2B) is referred to as a tetramer, half-nucleosome or hemisome (Figure 4). Work done in human cells resulted

in similar observations regarding the equimolar presence of core histones with particle heights and volumes fitting well with the half-nucleosome model as compared to H3 nucleosomes (Dimitriadis et al., 2010).

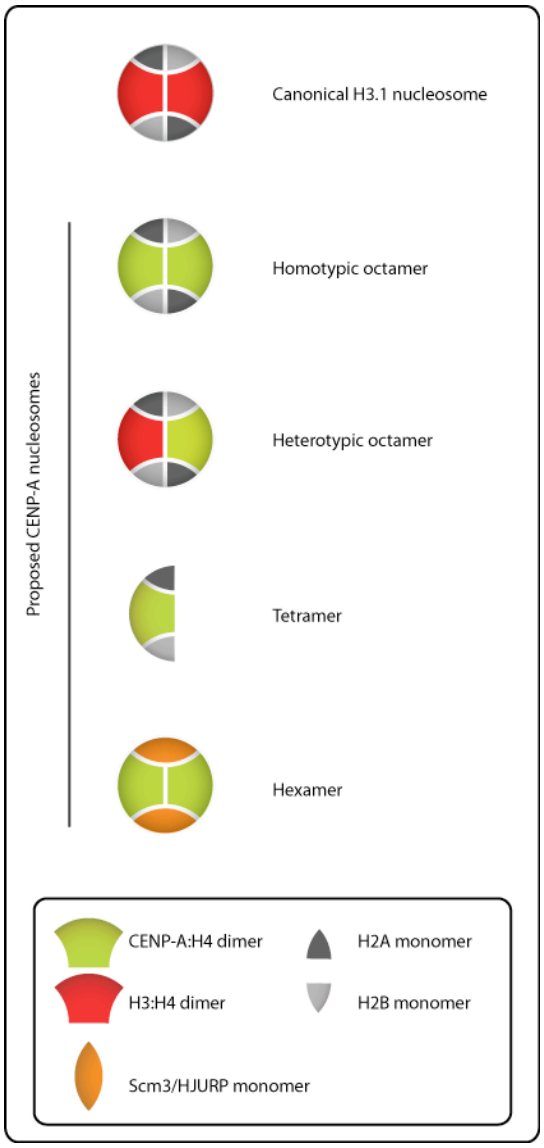

Figure 4: Different models for the composition of CENP-A nucleosomes.

These models differ in size, DNA wrapping and number as well as identity of their components. For a more detailed description of the differences between each model, please see the text.

Immuno-**e**lectron **m**icroscopy (Immuno-EM) data also suggest that the quantity of histones within each CENP-A nucleosomal particle matches with the half-nucleosome structure (Dimitriadis et al., 2010).

Recently, it was reported that the budding yeast centromeres are composed of a single Cse4-containing nucleosome wrapping about 80 bp of DNA, half the length of a canonical nucleosome, once in a right-handed manner (Henikoff and Henikoff, 2012; Krassovsky et al., 2012). The right-handedness of DNA wrapping in centromeric nucleosomes has been reported by other studies as well (Furuyama and Henikoff, 2009; Huang et al., 2011). ChIP data also demonstrated the occupancy of H2A at these sites all together consistent with the existence of a Cse4 hemisome at the centromere (Krassovsky et al., 2012). In support of this, the same study that found Cse4 nucleosomes (octamers) in chromosome arms reported tetramers in centromeres (Camahort et al., 2009). It should be noted that *in vivo* calibrated fluorescence intensity measurements of GFP:Cse4 are not consistent with a single copy of Cse4 at each centromere (Coffman et al., 2011; Lawrimore et al., 2011). Thus it is entirely possible that the conditions used for purification induce a hemisome like artifact and that does not exist or is not stable *in vivo*.

Intriguingly, the observations presented above about centromeric nucleosomes, are reminiscent of an early model suggesting that octameric nucleosomes are in fact constituted from symmetrical half-nucleosome pairs capable of independent existence (Weintraub et al., 1976). This was supported by work on SV40 minichromosome which primarily consists of about 20-25 nucleosomes as shown by electron microscopy (EM). Incubation of purified minichromosomes at low ionic strengths was however reported to induce the doubling of the number of beads-on-a-string, reduction of the dimensions of resulting particles and more interestingly longer inter-particle distances, all suggesting the splitting of octameric nucleosomes into half-nucleosomes (Lavelle and Prunell, 2007; Oudet et al., 1978). Similar

observations were made with cellular chromatin (Oudet et al., 1978). However, the occurrence of this conversion under cellular conditions remains to be shown to date.

On the other hand, the observation that nucleosomes can be found in various conformational states (Lavelle and Prunell, 2007), examples of which include Archaeal nucleosomes consisting solely of $(H3:H4)_2$ tetramers (Reeve et al., 1997), eukaryotic reversomes generated upon depletion of H2A:H2B dimers with right-handed DNA wrapping (Lavelle and Prunell, 2007) and the more recent proposed heterotetramer formation of CENPs-T:W:S:X (Nishino et al., 2012) capable of supercoiling DNA similar to nucleosomes, supports the possibility of tetrameric half-nucleosomes (hemisomes) residing in certain regions of the genome such as the centromere.

Other proposed forms of CENP-A nucleosomes:

Using a modified sequential immuno-precipitation technique in budding yeast, H3 was recently reported to co-occupy the centromeric DNA along with Cse4 and other core histones in a cell cycle independent manner (Lochmann and Ivanov, 2012) suggesting the potential existence of $(Cse4:H4)(H3:H4)(H2A:H2B)_2$ heterotypic octamers (Figure 4). However, it is not clear if stable association of H3 with Cse4 containing nucleosome is in the form of a heterotypic octamer or non-nucleosomal associations.

Additionally, a $(Cse4:H4)_2(Scm3)_2$ hexameric organization has also been proposed for the centromeric chromatin in budding yeast (Figure 4) (Mizuguchi et al., 2007). However, a number of key observations soon detracted support for stable occurrence of this structure in centromeric chromatin. These include the previously mentioned structural barriers occluding Cse4 and H4 interaction with DNA (Zhou et al., 2011) and the fact that over-expression of

Cse4 in an Scm3Δ background, can rescue the Scm3 null phenotype (Camahort et al., 2009) suggesting that Scm3 is dispensable for centromere organization.

The tetramer to octamer transition model; towards a dispute settlement?

The controversial observations regarding the nature of CENP-A nucleosomes possibly stem from different chromatin preparation techniques, the stabilization of transient intermediates or the co-existence of more than one CENP-A nucleosome type under certain conditions. No matter the technical difference, a potential structural dynamics model for CENP-A containing nucleosomes through the cell cycle would be an important step forward in understanding centromere biology.

In this regard, an octamer to tetramer conversion model had been previously proposed based on which octameric CENP-A nucleosomes are split into tetrameric half-nucleosomes upon the passage of the replication fork in S phase allowing the equal inheritance of the epigenetic mark to the daughter strands (Figure 5A) (Allshire and Karpen, 2008; Probst et al., 2009). The resultant tetramers were proposed to be maintained throughout G2/M but converted into octamers in G1 following incorporation of new CENP-A by HJURP. This model, while providing a possible mechanism for the preservation of centromeric identity, had never been experimentally validated.

Interestingly, two recent studies co-published in *Cell* (Bui et al., 2012; Shivaraju et al., 2012) provide evidence for a novel cell cycle-coupled structural transition of CENP-A nucleosomes in human cells and budding yeast (Figure 5B). AFM-based analysis of immuno-precipitated CENP-A nucleosomes from cell cycle staged human cells revealed that centromeric nucleosomes changed in size depending on cell cycle timing.

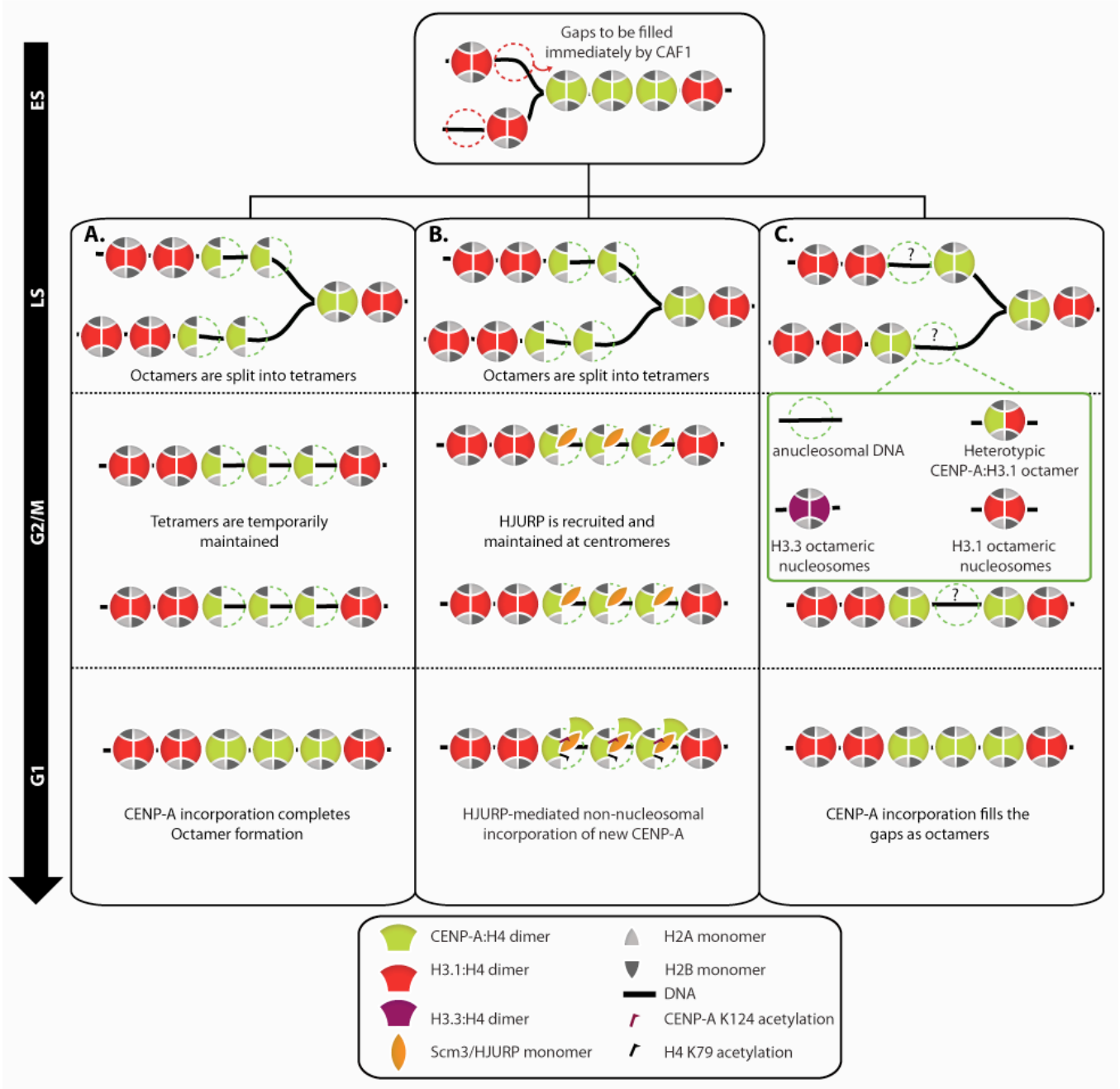

Figure 5: Proposed structural dynamics of centromeric nucleosomes throughout the cell cycle.

A. An octamer to tetramer transition model in which passage of the replication fork splits the pre-existing octameric CENP-A nucleosomes into tetramers allowing equal inheritance of the epigenetic mark to daughter strands. In this model, HJURP exclusively is found at the centromere in G1 and mediates the reconstruction of octamers by incorporation of new CENP-A. B. The tetramer to octamer transition model. CENP-A:H4:HJURP complex is recruited to the centromeric chromatin in G1 and is assumed to associate to the pre-exisiting CENP-A tetramers extra-nucleosomally. Post-translational modifications (PTMs) of tetrameric CENP-A nucleosomes along with the presence of HJURP impede the stable incorporation of new CENP-A:H4 dimers into nucleosomes. In late G1/early S, action of certain protein chaperones and remodeling complexes is proposed to facilitate the tetramer to octamer transition by affecting the PTMs of histones and chromatin structure resulting in tetramer to octamer transition. Likewise to the previous model, passage of the replication fork splits the octamers to tetramers. The tetramers are further stabilized by the reassociation of HJURP in G2 and presumed to be maintained in mitosis to form the kinetochore plate. **C.** In an alternative model, where CENP-A exists in octameric nuclesomes throughout the cell cycle, gaps generated as a result of the passage of the replication fork will be either maintained or could be transiently filled with multiple possible placeholder structures. In this model, HJURP is also found at the centromere exclusively in G1 to participate in new CENP-A assembly.

Tetrameric dimensions of CENP-A nucleosomes in G1 were reported to undergo a transition to octameric dimensions during S phase and revert back to tetrameric state in G2 and maintained through mitosis (Bui et al., 2012). Intriguingly, the authors also report cyclic association, dissociation and re-association of HJURP to the centromeric chromatin in G1, S and G2 phases respectively (Figure 5B). Purification of DNA-bound CENP-A:H4 from G1/S arrested cells followed by mass-spectroscopic analysis identified two previously unknown covalent modifications: acetylation of CENP-A K124 and H4 K79. These modifications and the presence of HJURP were proposed to prevent stable octamer formation in G1/S. However, as the cell enters the S phase, the authors speculated that opening of centromeric chromatin would be concomitant with the resolution of these modifications. This biochemical change could be coupled to the release of HJURP physically allowing the action of chromatin remodelers to trigger the generation octamers. The dynamics of HJURP at the centromere and the reformation of tetramers in G2 might indicate the role of HJURP in this reversal transition following DNA replication. However, HJURP in human cells has been previously shown to localize to centromeres exclusively during CENP-A loading in G1(Dunleavy et al., 2009; Foltz et al., 2009) and the G2 reappearance of HJURP at the centromeres has never been reported by other groups. Given the fact that nascent CENP-A is chaperoned by HJURP after synthesis in G2, cytoplasmic contamination of chromatin lysate could be a potential source for the detection of HJURP in G2/M CENP-A pull down. On the other hand, the authors report that HJURP is absent from S phase CENP-A pull down. If the possibility of cytoplasmic contamination is true, this would nicely explain the absence of HJURP in S phase CENP-A pull down since HJURP has already completed deposition of its cargo and thus even if the

chromatin prep does contain cytoplasmic contamination, HJURP will not copurify with CENP-A any longer.

Work in *Saccharomyces cerevisiae* and *Candida albicans* using fluorescence correlation spectroscopy (FCS) also suggest the presence of a single copy of Cse4 at each centromere (Shivaraju et al., 2012). This seems to be the case for G1, S, G2 and metaphase, however during anaphase B the authors note a "doubling" of Cse4 and speculate this is due to a tetramer to octamer transition. FCS can very accurately measure protein complexes in living cells by recording peak intensities of diffusing molecules and comparing these over varying time scales (Bulseco and Wolf, 2007) (usually up to 2 minutes), detecting auto-correlations. A major caveat with this technique is that while extremely accurate for freely diffusing complexes, it is less accurate for slow diffusing structures (Krichevsky and Bonnet, 2002) (in this case mitotic centromeres bound to microtubules). This is in part due to the fact that FCS is based on peak intensity of a diffraction limited spot. Thus, slowly diffusing or slightly dispersed (greater than the measurement spot) structures such as centromeres will not all be measured in the time scale required.

Interestingly, centromeres are much less dispersed in anaphase (compared to metaphase) (Pearson et al., 2001), which could allow for more accurate measurements and thus explain the difference reported. Nonetheless, using FRET and sequential ChIPs the authors show that Cse4:Cse4 interaction does indeed take place increasingly in anaphase B (Shivaraju et al., 2012). The tetramer to octamer transition is also concomitant with the transient disappearance of Scm3 from centromeres in a short time window corresponding to anaphase B. The mutually exclusive relationship between presence of Scm3/HJURP and the Cse4/CENP-A dimers could be attributed to the unique HFD of CENP-A harboring a shared interface for interaction with Scm3/HJURP or another molecule of CENP-A.

Even though the generation of differential CENP-A nucleosome types and the transition mechanisms remain largely unknown, it is presumable that such a dynamic behavior may require a tight regulation for timing and the concerted action of chromatin remodeling factors. It would be interesting to investigate the occurrence of K124- and K79-like modifications in yeast Cse4 and H4. In addition, the previous detection of CID-nucleosomes corresponding in dimensions to tetramers in interphase but octamers in mitotic *Drosophila* cells (Dalal et al., 2007) might reflect a similar cell-cycle regulated transition formerly unexplored.

Implications of the structure of CENP-A nucleosomes: why does it matter after all?

Emerging evidence suggests that the assembly of CENP-A and thus propagation of the epigenetic mark occurs through three major steps: licensing by KNL-2/M18BP1 (Fujita et al., 2007; Maddox et al., 2007) and Mis18 (Hayashi et al., 2004), incorporation via HJURP (Bernad et al., 2011; Dunleavy et al., 2009; Foltz et al., 2009) and maintenance by

MgcRacGAP (Lagana et al., 2010). During S phase, as the replication fork forges ahead, the pre-existing population of CENP-A nucleosomes is halved and thereby inherited to daughter strands to preserve centromeric identity (Allshire and Karpen, 2008). The dilution of CENP-A nucleosomes and the replication-independent incorporation of CENP-A raise the possibility of the formation of various placeholder structures in S phase (Figure 5C). In the case of human cells, according to recent data, the CENP-A nucleosomal population is proposed to undergo the tetramer to octamer transition in front of the replication fork approaching the centromeric DNA (Bui et al., 2012). Figure 6 depicts possible steps in this transition.

It is not clear what exactly signals the speculated reversion of octamers into tetramers at the end of S phase. Passage of the fork may, via unknown mechanisms or interactions, trigger not only the splitting of the pre-existing CENP-A nucleosomes, but also the reversion of octamers into tetramers. This may in turn coincide with the reappearance of HJURP at the end of S phase in human cells. In contrast to canonical nucleosomes, incorporation of newly synthesized CENP-A nucleosomes does not accompany DNA replication (Jansen et al., 2007). CENP-A assembly in the mammalian system requires exit from mitosis (Jansen et al., 2007) and takes place during late M/G1 phase of the cell cycle in the mammalian and embryonic *Drosophila* systems (Mellone et al., 2011; Schuh et al., 2007). However, the mechanism of specific recognition of the centromere by CENP-A assembly proteins remains largely unknown. It is assumed that these proteins might recognize a specialized chromatin structure, certain contact sites on CENP-A-containing nucleosomes or a non-conventional nucleosome form exclusively found at the centromeric chromatin. Given the proposed atypical CENP-A nucleosomes, regardless of the model, an entertaining speculation would be that the heteroclite structure of CENP-A nucleosomes might provide the green light for CENP-A assembly

machinery to repopulate the centromere in preparation for the subsequent mitosis. In addition, these atypical structures may serve as recognition sites for kinetochore assembly during mitosis.

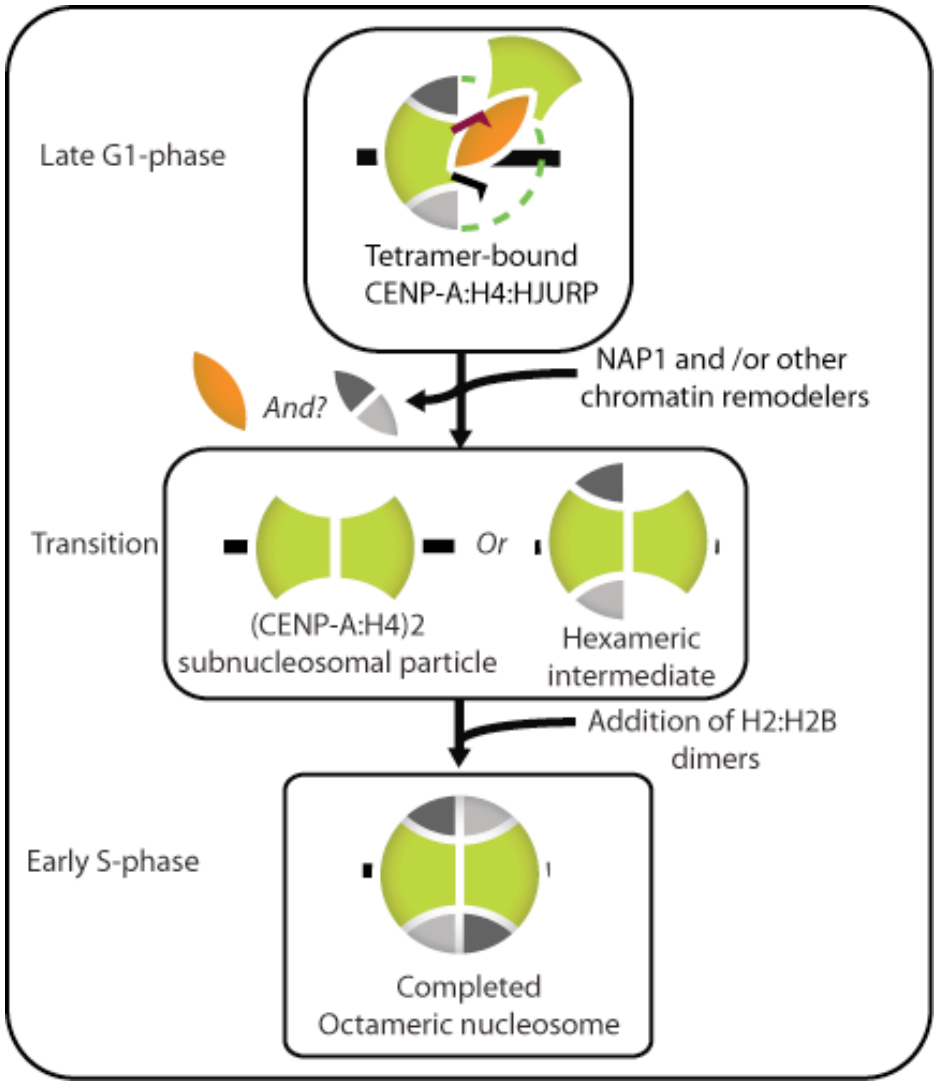

Figure 6 : Possible steps of the tetramer to octamer transition.

Upon transition of the cell to S phase, HJURP is released from the centromeric chromatin and certain posttranslational modifications and of CENP-A tetramers are resolved by yet-to-be-identified factors. In the first scenario, H2A:H2B dimers are also temporarily removed

allowing the tetrameric (CENP-A:H4)$_2$ complex to form and interact with DNA. This reaction is speculated to be mediated by NAP1, which is a chromatin remodeler. Addition of a pair of H2A:H2B dimers will complete the octamer formation. In the second scenario, however, incorporation of new CENP-A:H4 as well as H2B:H2B dimers does not require disassembly of the pre-existing CENP-A tetramers.

Conclusion:

This review summarizes major features of CENP-A-containing complexes on sub-nucleosomal, nucleosomal and chromatin levels. The debate over the true molecular nature of the CENP-A epigenetic mark remains to be resolved as many questions are still unanswered. For example, what could be the biological significance of the tetramer to octamer transition in S phase for mammalian cells or anaphase B in case of yeast? What are the factors and mechanisms involved? In the coming exciting years of research, high-resolution imaging and biochemical approaches hold promise to pave the way for answering these questions.

Acknowledgment:

We are grateful to all members of Paul and Amy Maddox labs. Special thanks to Joel Ryan for his valuable assistance with the figures. P.S.M. is the Canada Research Chair in Cell Division and Chromosomal Organization and supported by research grants from the CIHR (MOP-106548) and CCSRI (700824).

References:

Allshire, R.C., and Karpen, G.H. (2008). Epigenetic regulation of centromeric chromatin: old dogs, new tricks? Nature reviews Genetics *9*, 923-937.

Barnhart, M.C., Kuich, P.H., Stellfox, M.E., Ward, J.A., Bassett, E.A., Black, B.E., and Foltz, D.R. (2011). HJURP is a CENP-A chromatin assembly factor sufficient to form a functional de novo kinetochore. The Journal of cell biology *194*, 229-243.

Bassett, E.A., DeNizio, J., Barnhart-Dailey, M.C., Panchenko, T., Sekulic, N., Rogers, D.J., Foltz, D.R., and Black, B.E. (2012). HJURP uses distinct CENP-A surfaces to recognize and to stabilize CENP-A/histone H4 for centromere assembly. Developmental cell *22*, 749-762.

Bernad, R., Sanchez, P., Rivera, T., Rodriguez-Corsino, M., Boyarchuk, E., Vassias, I., Ray-Gallet, D., Arnaoutov, A., Dasso, M., Almouzni, G.*, et al.* (2011). Xenopus HJURP and condensin II are required for CENP-A assembly. The Journal of cell biology *192*, 569-582.

Black, B.E., Brock, M.A., Bedard, S., Woods, V.L., Jr., and Cleveland, D.W. (2007a). An epigenetic mark generated by the incorporation of CENP-A into centromeric nucleosomes. Proceedings of the National Academy of Sciences of the United States of America *104*, 5008-5013.

Black, B.E., Foltz, D.R., Chakravarthy, S., Luger, K., Woods, V.L., Jr., and Cleveland, D.W. (2004). Structural determinants for generating centromeric chromatin. Nature *430*, 578-582.

Black, B.E., Jansen, L.E., Maddox, P.S., Foltz, D.R., Desai, A.B., Shah, J.V., and Cleveland, D.W. (2007b). Centromere identity maintained by nucleosomes assembled with histone H3 containing the CENP-A targeting domain. Molecular cell *25*, 309-322.

Bui, M., Dimitriadis, E.K., Hoischen, C., An, E., Quenet, D., Giebe, S., Nita-Lazar, A., Diekmann, S., and Dalal, Y. (2012). Cell-Cycle-Dependent Structural Transitions in the Human CENP-A Nucleosome In Vivo. Cell *150*, 317-326.

Bulseco, D.A., and Wolf, D.E. (2007). Fluorescence correlation spectroscopy: molecular complexing in solution and in living cells. Methods in cell biology *81*, 525-559.

Camahort, R., Shivaraju, M., Mattingly, M., Li, B., Nakanishi, S., Zhu, D., Shilatifard, A., Workman, J.L., and Gerton, J.L. (2009). Cse4 is part of an octameric nucleosome in budding yeast. Molecular cell *35*, 794-805.

Cho, U.S., and Harrison, S.C. (2011). Recognition of the centromere-specific histone Cse4 by the chaperone Scm3. Proceedings of the National Academy of Sciences of the United States of America *108*, 9367-9371.

Cleveland, D.W., Mao, Y., and Sullivan, K.F. (2003). Centromeres and kinetochores: from epigenetics to mitotic checkpoint signaling. Cell *112*, 407-421.

Coffman, V.C., Wu, P., Parthun, M.R., and Wu, J.Q. (2011). CENP-A exceeds microtubule attachment sites in centromere clusters of both budding and fission yeast. The Journal of cell biology *195*, 563-572.

Conde e Silva, N., Black, B.E., Sivolob, A., Filipski, J., Cleveland, D.W., and Prunell, A. (2007). CENP-A-containing nucleosomes: easier disassembly versus exclusive centromeric localization. Journal of molecular biology *370*, 555-573.

Dalal, Y., Wang, H., Lindsay, S., and Henikoff, S. (2007). Tetrameric structure of centromeric nucleosomes in interphase Drosophila cells. PLoS Biol *5*, e218.

Dimitriadis, E.K., Weber, C., Gill, R.K., Diekmann, S., and Dalal, Y. (2010). Tetrameric organization of vertebrate centromeric nucleosomes. Proceedings of the National Academy of Sciences of the United States of America *107*, 20317-20322.

Dunleavy, E.M., Roche, D., Tagami, H., Lacoste, N., Ray-Gallet, D., Nakamura, Y., Daigo, Y., Nakatani, Y., and Almouzni-Pettinotti, G. (2009). HJURP is a cell-cycle-dependent maintenance and deposition factor of CENP-A at centromeres. Cell *137*, 485-497.

Foltz, D.R., Jansen, L.E., Bailey, A.O., Yates, J.R., 3rd, Bassett, E.A., Wood, S., Black, B.E., and Cleveland, D.W. (2009). Centromere-specific assembly of CENP-a nucleosomes is mediated by HJURP. Cell *137*, 472-484.

Fujita, Y., Hayashi, T., Kiyomitsu, T., Toyoda, Y., Kokubu, A., Obuse, C., and Yanagida, M. (2007). Priming of centromere for CENP-A recruitment by human hMis18alpha, hMis18beta, and M18BP1. Developmental cell *12*, 17-30.

Fukagawa, T. (2004). Centromere DNA, proteins and kinetochore assembly in vertebrate cells. Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology *12*, 557-567.

Furuyama, T., and Henikoff, S. (2009). Centromeric nucleosomes induce positive DNA supercoils. Cell *138*, 104-113.

Guse, A., Carroll, C.W., Moree, B., Fuller, C.J., and Straight, A.F. (2011). In vitro centromere and kinetochore assembly on defined chromatin templates. Nature *477*, 354-358.

Hayashi, T., Fujita, Y., Iwasaki, O., Adachi, Y., Takahashi, K., and Yanagida, M. (2004). Mis16 and Mis18 are required for CENP-A loading and histone deacetylation at centromeres. Cell *118*, 715-729.

Henikoff, S., and Henikoff, J.G. (2012). "Point" centromeres of Saccharomyces harbor single centromere-specific nucleosomes. Genetics *190*, 1575-1577.

Howman, E.V., Fowler, K.J., Newson, A.J., Redward, S., MacDonald, A.C., Kalitsis, P., and Choo, K.H. (2000). Early disruption of centromeric chromatin organization in centromere protein A (Cenpa) null mice. Proceedings of the National Academy of Sciences of the United States of America *97*, 1148-1153.

Hu, H., Liu, Y., Wang, M., Fang, J., Huang, H., Yang, N., Li, Y., Wang, J., Yao, X., Shi, Y.*, et al.* (2011). Structure of a CENP-A-histone H4 heterodimer in complex with chaperone HJURP. Genes & development *25*, 901-906.

Huang, C.C., Chang, K.M., Cui, H., and Jayaram, M. (2011). Histone H3-variant Cse4-induced positive DNA supercoiling in the yeast plasmid has implications for a plasmid origin of a chromosome centromere. Proceedings of the National Academy of Sciences of the United States of America *108*, 13671-13676.

Jansen, L.E., Black, B.E., Foltz, D.R., and Cleveland, D.W. (2007). Propagation of centromeric chromatin requires exit from mitosis. The Journal of cell biology *176*, 795-805.

Krassovsky, K., Henikoff, J.G., and Henikoff, S. (2012). Tripartite organization of centromeric chromatin in budding yeast. Proceedings of the National Academy of Sciences of the United States of America *109*, 243-248.

Krichevsky, O., and Bonnet, G. (2002). Fluorescence correlation spectroscopy: the technique and its applications. Rep Prog Phys *65*, 251-297.

Lagana, A., Dorn, J.F., De Rop, V., Ladouceur, A.M., Maddox, A.S., and Maddox, P.S. (2010). A small GTPase molecular switch regulates epigenetic centromere maintenance by stabilizing newly incorporated CENP-A. Nature cell biology *12*, 1186-1193.

Lavelle, C., and Prunell, A. (2007). Chromatin polymorphism and the nucleosome superfamily: a genealogy. Cell cycle *6*, 2113-2119.

Lawrimore, J., Bloom, K.S., and Salmon, E.D. (2011). Point centromeres contain more than a single centromere-specific Cse4 (CENP-A) nucleosome. The Journal of cell biology *195*, 573-582.

Lochmann, B., and Ivanov, D. (2012). Histone H3 localizes to the centromeric DNA in budding yeast. PLoS genetics *8*, e1002739.

Maddox, P.S., Hyndman, F., Monen, J., Oegema, K., and Desai, A. (2007). Functional genomics identifies a Myb domain-containing protein family required for assembly of CENP-A chromatin. The Journal of cell biology *176*, 757-763.

Mellone, B.G., Grive, K.J., Shteyn, V., Bowers, S.R., Oderberg, I., and Karpen, G.H. (2011). Assembly of Drosophila centromeric chromatin proteins during mitosis. PLoS genetics *7*, e1002068.

Nishino, T., Takeuchi, K., Gascoigne, K.E., Suzuki, A., Hori, T., Oyama, T., Morikawa, K., Cheeseman, I.M., and Fukagawa, T. (2012). CENP-T-W-S-X forms a unique centromeric chromatin structure with a histone-like fold. Cell *148*, 487-501.

Oudet, P., Germond, J.E., Bellard, M., Spadafora, C., and Chambon, P. (1978). Nucleosome structure. Philosophical transactions of the Royal Society of London Series B, Biological sciences *283*, 241-258.

Palmer, D.K., O'Day, K., Trong, H.L., Charbonneau, H., and Margolis, R.L. (1991). Purification of the centromere-specific protein CENP-A and demonstration that it is a distinctive histone. Proceedings of the National Academy of Sciences of the United States of America *88*, 3734-3738.

Panchenko, T., Sorensen, T.C., Woodcock, C.L., Kan, Z.Y., Wood, S., Resch, M.G., Luger, K., Englander, S.W., Hansen, J.C., and Black, B.E. (2011). Replacement of histone H3 with CENP-A directs global nucleosome array condensation and loosening of nucleosome superhelical termini. Proceedings of the National Academy of Sciences of the United States of America *108*, 16588-16593.

Pearson, C.G., Maddox, P.S., Salmon, E.D., and Bloom, K. (2001). Budding yeast chromosome structure and dynamics during mitosis. The Journal of cell biology *152*, 1255-1266.

Probst, A.V., Dunleavy, E., and Almouzni, G. (2009). Epigenetic inheritance during the cell cycle. Nature reviews Molecular cell biology *10*, 192-206.

Reeve, J.N., Sandman, K., and Daniels, C.J. (1997). Archaeal histones, nucleosomes, and transcription initiation. Cell *89*, 999-1002.

Schuh, M., Lehner, C.F., and Heidmann, S. (2007). Incorporation of Drosophila CID/CENP-A and CENP-C into centromeres during early embryonic anaphase. Current biology : CB *17*, 237-243.

Sekulic, N., Bassett, E.A., Rogers, D.J., and Black, B.E. (2010). The structure of (CENP-A-H4)(2) reveals physical features that mark centromeres. Nature *467*, 347-351.

Shelby, R.D., Monier, K., and Sullivan, K.F. (2000). Chromatin assembly at kinetochores is uncoupled from DNA replication. The Journal of cell biology *151*, 1113-1118.

Shivaraju, M., Unruh, J.R., Slaughter, B.D., Mattingly, M., Berman, J., and Gerton, J.L. (2012). Cell-cycle-coupled structural oscillation of centromeric nucleosomes in yeast. Cell *150*, 304-316.

Shuaib, M., Ouararhni, K., Dimitrov, S., and Hamiche, A. (2010). HJURP binds CENP-A via a highly conserved N-terminal domain and mediates its deposition at centromeres. Proceedings of the National Academy of Sciences of the United States of America *107*, 1349-1354.

Sullivan, K.F., Hechenberger, M., and Masri, K. (1994). Human CENP-A contains a histone H3 related histone fold domain that is required for targeting to the centromere. J Cell Biol *127*, 581-592.

Tachiwana, H., Kagawa, W., Shiga, T., Osakabe, A., Miya, Y., Saito, K., Hayashi-Takanaka, Y., Oda, T., Sato, M., Park, S.Y.*, et al.* (2011). Crystal structure of the human centromeric nucleosome containing CENP-A. Nature *476*, 232-235.

Tomonaga, T., Matsushita, K., Yamaguchi, S., Oohashi, T., Shimada, H., Ochiai, T., Yoda, K., and Nomura, F. (2003). Overexpression and mistargeting of centromere protein-A in human primary colorectal cancer. Cancer Res *63*, 3511-3516.

Vagnarelli, P., Ribeiro, S.A., and Earnshaw, W.C. (2008). Centromeres: old tales and new tools. FEBS letters *582*, 1950-1959.

Weintraub, H., Worcel, A., and Alberts, B. (1976). A model for chromatin based upon two symmetrically paired half-nucleosomes. Cell *9*, 409-417.

Yoda, K., Ando, S., Morishita, S., Houmura, K., Hashimoto, K., Takeyasu, K., and Okazaki, T. (2000). Human centromere protein A (CENP-A) can replace histone H3 in nucleosome reconstitution in vitro. Proceedings of the National Academy of Sciences of the United States of America *97*, 7266-7271.

Zhang, W., Colmenares, S.U., and Karpen, G.H. (2012). Assembly of Drosophila centromeric nucleosomes requires CID dimerization. Molecular cell *45*, 263-269.

Zhou, Z., Feng, H., Zhou, B.R., Ghirlando, R., Hu, K., Zwolak, A., Miller Jenkins, L.M., Xiao, H., Tjandra, N., Wu, C.*, et al.* (2011). Structural basis for recognition of centromere histone variant CenH3 by the chaperone Scm3. Nature *472*, 234-237.

Chapter 2

Octameric CENP-A nucleosomes are present at human centromeres throughout the cell cycle

Abbas Padeganeh[1, 3], Joël Ryan[1ψ], Jacques Boisvert[1,4ψ], Anne-Marie Ladouceur[1, 3], Jonas F. Dorn[1] and Paul S. Maddox[1,2 *]

[1]Institute for Research in Immunology and Cancer (IRIC)

[2]Department of Pathology and Cell Biology,

[3]Graduate programs in Molecular Biology-Systems Biology Option,

[4]Graduate programs in Bioinformatics,

Faculty of Medicine, University of Montreal, H3T 1J4, Quebec, Canada

[ψ] These authors contributed equally to this work

[*]To whom correspondence should be addressed

Institute for Research in Immunology and Cancer (IRIC), Dept of Pathology and Cell Biology,

Université de Montréal

P.O. Box 6128, Station Centre-Ville

Montréal QC, H3C 3J7

CANADA

In this chapter, I present the development of a high-resolution single-molecule imaging assay to address the molecular entity of the centromeric chromatin epigenetic mark and its compositional dynamics during the cell cycle.

The material presented herein was published in ***Current Biology. 2013 May 6;23(9):764-9*** and featured in a News & Views article in *Nat Struct Mol Biol*. Jun 2013;20(6):648-50)

Author contribution:

PSM and AP designed the project. AP performed the wet lab experiments, wrote the manuscript and prepared the images. JR and AML assisted with the experiments. JB and JFD developed the automated data analysis software. PSM and AP and JR analyzed the data and PSM edited the manuscript.

Summary:

The presence of a single centromere on each chromosome that signals formation of a mitotic kinetochore is central to accurate chromosome segregation (Cleveland et al., 2003). The histone H3 variant CENP-A is critical for centromere identity and function; CENP-A chromatin acts as an epigenetic mark to direct both centromere and kinetochore assembly (Allshire and Karpen, 2008; Black et al., 2007a; De Rop et al., 2012). Interpreting the centromere epigenetic mark ensures propagation of a single centromere per chromosome to maintain ploidy. Thus, understanding the nature of CENP-A chromatin is crucial for all cell divisions. However there are ongoing debates over the fundamental composition of centromeric chromatin. Here we show that natively assembled human CENP-A nucleosomes are octameric throughout the cell cycle. Using TIRF-coupled photobleaching-assisted copy number counting of single nucleosomes obtained from cultured cells, we find that the majority of CENP-A nucleosomes contain CENP-A dimers. In addition, we detect the presence of H2B and H4 in these nucleosomes. Surprisingly, CENP-A associated with the chaperone HJURP can exist as either monomer or dimer, indicating possible assembly intermediates. Thus, our findings indicate that octameric CENP-A nucleosomes mark the centromeric region to ensure proper epigenetic inheritance and kinetochore assembly.

Highlights

1) CENP-A octameric nucleosomes epigenetically mark centromeres

2) CENP-A associated with the chaperone HJURP can exist as either monomer or dimer

Results and Discussion:

Conservation of ploidy requires inheritance of an equal number of chromosomes during each cell division. Central to this is the presence of a single kinetochore on each chromatid. Kinetochores are protein super-structures assembled during mitosis at centromere regions of chromosomes to link sister chromatids to microtubules emanating from opposite spindle poles. Consequently, centromere singularity generates accurate cell division (Cleveland et al., 2003). The defining feature of centromeres is understood to be the presence of a histone H3 variant referred to as Centromere Protein-A (CENP-A, also called CenH3) in centromeric nucleosomes (Black et al., 2007a; De Rop et al., 2012). CENP-A is proposed to mark centromeres epigenetically (Allshire and Karpen, 2008; Black et al., 2007a) as evidenced by rarely documented cases of neocentromeres, where CENP-A chromatin forms a functional centromere on an ectopic chromosomal region distinct from the original genomic locus. In sum, CENP-A-containing chromatin encodes the epigenetic information recognized, read and interpreted by proteins required for centromere (Foltz et al., 2009; Fujita et al., 2007; Lagana et al., 2010; Shuaib et al., 2010) and kinetochore assembly (Cleveland et al., 2003). Thus, understanding the biochemical nature of CENP-A chromatin is critical for understanding both epigenetic centromere propagation and mitotic chromosome segregation.

The majority of CENP-A containing complexes in cells contain two CENP-A molecules:

Unlike other epigenetic codes (e.g. histone methylation patterns regulating promoter activity), the centromere epigenetic code is as yet largely unknown. Extensive efforts focused to determine the structure of CENP-A-containing nucleosomes have led to contradicting observations. The culmination is a series of hypotheses each differing from one another in the

species and number of centromere chromatin core components (Padeganeh et al., 2013). These include 1) the octamer model (Panchenko et al., 2011; Tachiwana et al., 2011; Tachiwana and Kurumizaka, 2011; Zhang et al., 2012) supported mainly by *in vitro* nucleosome reconstitution experiments proposing a $(CENP\text{-}A:H4)_2$ $(H2A:H2B)_2$ composition, 2) the tetramer model (Dalal et al., 2007; Dimitriadis et al., 2010; Krassovsky et al., 2012) stemming from Atomic Force Microscopy data on immunoprecipitated CENP-A chromatin proposing a (CENP-A:H4)(H2A:H2B) composition and 3) the hexamer model (Mizuguchi et al., 2007) proposing a hybrid of CENP-A:H4 and the chaperone protein HJURP (also called Scm3, hereafter HJURP).

To determine the molecular composition of native CENP-A nucleosomes, we extracted single nucleosome core particles from clonal HeLa cells stably expressing a YFP fusion to CENP-A (see below and methods). For our analysis, we assumed that CENP-A-YFP and unlabeled CENP-A have equal probability of chromatin incorporation, and corrected for the incomplete labeling accordingly (see below and methods). In order to assess the stoichiometry of CENP-A in centromeric nucleosomes, we employed photo-bleaching-assisted copy number counting (PA-CNC, a modified version of SiM-Pull (Jain et al., 2011)). We assembled a homemade flow chamber and functionalized it with YFP-nanobodies (Rothbauer et al., 2008). This surface would act as a nanotrap for CENP-A-YFP containing complexes, effectively immunoprecipitating YFP containing particles from the whole chromatin extract (Figure 1A). We evaluated the specificity of our assay using control (no-YFP) HeLa chromatin lysates with and without YFP nanobodies and found minimal contamination, indicating that the analyzed signals were YFP derived (Figure 1B, and Supplement 1A). Importantly, nanobodies are monoclonal and single chained, therefore each nanobody isolates one and only one YFP

containing complex. The final chamber, when visualized by Total Internal Reflection Fluorescence (TIRF) microscopy, results in single isolated diffraction limited spots each representing a single CENP-A-YFP containing chromatin particle (Figure 1B). Thus, this system allows direct visual analysis of CENP-A-YFP stoichiometry in native assembled complexes. Our method has the additional benefit that there are very few manipulations of the sample as opposed to conventional fractionation techniques, reducing potential artifacts. A caveat of our method is that the TIRF illumination field is uneven, a common issue with objective-based TIRF systems. Uneven illumination precludes comparison of absolute intensities between individual complexes or experiments, however does not affect relative measurements of individual complexes.

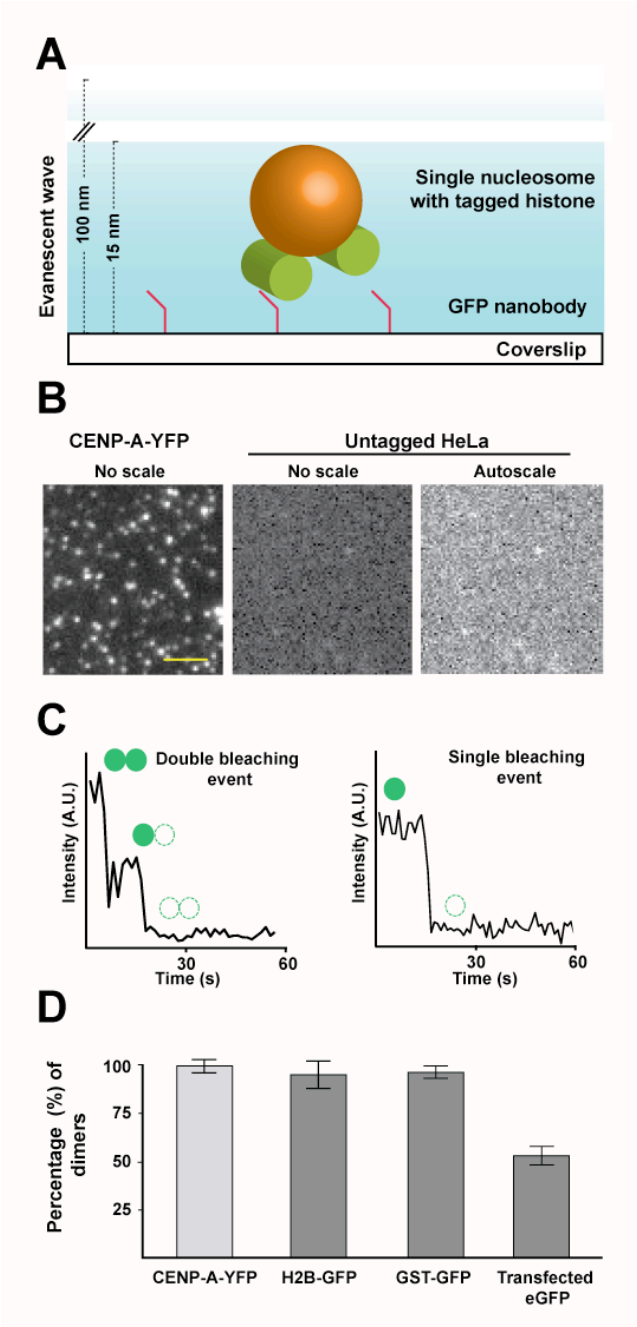

Figure 1: Asynchronous CENP-A-YFP expressing cells contain dimers of CENP-A at centromeric nucleosomes.

A. Schematic representation of our PA-CNC set up. Using anti-YFP single-chained nanobodies allows the isolation of single CENP-A-YFP nucleosomes on the surface of a coverslip. Samples are allowed to photobleach completely by TIRF illumination. By plotting the fluorescence intensity fluctuations, detected for each diffraction limited spot, over time, the number of fluorophores in each spot is quantified by counting the number of discrete bleaching events. **B.** Representative TIRF images are shown for samples with and without addition of CENP-A-YFP nucleosomes. **C.** Example traces of pixel intensity fluctuation over time in double and single bleaching events. Camera integration time for all movies is 900 ms. (A.U. = arbitrary unit). **D.**The percentage of double bleaching events detected using our PA-CNC assay with purified nucleosomes from CENP-A-YFP and H2B-GFP stable cell lines, purified GST-GFP and cellular lysate from eGFP-transfected HeLa cells. The data presented herein are corrected for the expression levels of the labelled proteins and pre-bleaching where applicable. (scale bar = 20 μm; error bars indicate standard deviation).

A key feature differing in models for the structure of centromeric chromatin is the presence of one or two CENP-A molecules per nucleosome. This difference has obvious implications towards the mechanisms interpreting the centromere epigenetic mark and therefore chromosome segregation. To quantify the number of CENP-A-YFP molecules per individual complex detected in discrete diffraction limited spots, consecutive images were acquired while illuminating the sample until all fluorophores photobleached (Supplemental movie S1). Plotting the intensity of each diffraction limited spot as a function of time revealed minor fluctuations (random noise) overlaid on large drops in intensity (Figure 1C, and Supplement 2). Large drops were generated by permanent bleaching of individual YFP molecules, visible in graphs of intensity versus time. Photobleaching is stochastic and a direct measure of the number of YFP molecules per isolated complex. Because measuring intensity drops by hand is very labor intensive and subject to human error, we developed a custom software package (QUBE for QUantitative Bleaching Estimation, written in MATLAB) to automatically detect, segment, and measure spot intensities over time. This greatly increased the number of spots analyzed (thousands per condition) and yielded similar, however far more statistically significant, results to those obtained by manual analysis (data not shown). Periodically, large intensity increases following large drops were observed (Supplement 2, sample blinking event). These were likely due to the reported blinking behavior of YFP, or to the polarized nature of the TIRF illumination light. Regardless of the origin, blinking events are indicative of single molecules and therefore enhance our confidence in the singularity of analyzed particles. In fact, fluorescence blinking is the basis for single molecule super-resolution techniques such as STochastic Optical Reconstruction Microscopy (STORM) (Rust et al., 2006). Analysis of CENP-A-YFP nucleosomes isolated from asynchronous cells

suggested that the majority (99.0 ± 3.4%) of these nucleosomes contain two molecules of CENP-A-YFP (Figure 1D). Similar results were found for H2B-GFP (GFP crossreacts with the nanobody) nucleosomes (96.1 ± 7.1%), which are known to contain two copies of H2B (Figure 1D). Together, these data suggest that centromere nucleosomes contain two copies of CENP-A, consistent with canonical models.

To ensure our assay was probing the content of single nucleosomes, nucleosome core particles were isolated from total chromatin after nuclear lysis (to enrich for chromatin associated complexes) and subjected to micrococcal nuclease digestion. We confirmed that micrococcal nuclease digestion generated mono-nucleosomes by agarose gel electrophoresis, which revealed a single band near 150 bp, corresponding to the length of DNA wrapping individual nucleosomes (Figure 2A). Nuclease digestion has been used for decades to generate mono-nucleosomes, however it is unclear if this treatment may affect subnucleosomal complexes such as the proposed "hemisome". Importantly, in sonicated preparations not treated with nuclease, individual signals can be observed (at a much lower frequency) and these displayed a similar ratio of double bleaching events to that seen in nuclease treatment (data not shown) indicating that nuclease treatment did not significantly alter our ability to detect CENP-A monomers. An additional possible source of error is that our cell line expresses a fraction (90%) of fluorescently labeled CENP-A as determined by western blot (Figure 2D). As mentioned above, our analysis included correction for the ratio of YFP to endogenous protein (see methods and below). Our results also support the hypothesis that CENP-A nucleosomes are homo-octameric and not "hybrid" nucleosomes containing one copy of CENP-A and one copy of histone H3.

To further verify the accuracy of our assay, we counted fluorophores in control complexes with theoretically known stoichiometry. Bacterially purified GST-GFP dimers or cytosolic eGFP HeLa cell lysates were used to test the capacity of the assay to detect discrete bleaching events. As anticipated, we observed the major population of spots to contain double bleaching events in the GST-GFP sample (97.6 ± 3.2%) whereas the predominant population in the eGFP cell lysates showed a single bleaching event (Figure 1D). It is noteworthy that the observed percentage of double bleaching events in the cytosolic eGFP sample likely reflects intrinsic self-dimerization of the GFP molecule (Phillips, 1997). Additionally, analysis of H2B-GFP complexes that contained a different ratio of labeled to endogenous protein compared to our CENP-A-YFP samples (Figure 2D) confirmed that our corrections are accurate over a varying expression level. Thus, our assay reveals the number of fluorescent molecules per isolated complex.

CENP-A chromatin does not change stoichiometry over the cell cycle:

Centromere epigenetic repopulation occurs in G1 (De Rop et al., 2012). After mitosis, centromere levels of CENP-A double and then are halved during DNA replication, generating two daughter centromeres. Therefore the centromere epigenetic mark could be considered labile at two key points; G1 repopulation and S-phase distribution. Recently, two studies concluded that CENP-A stoichiometry would undergo a cell cycle transition (Bui et al., 2012; Shivaraju et al., 2012). Such a transition could be critically important for epigenetic recognition by downstream mechanisms regulating centromere propagation and/or kinetochore assembly. Thus, a possible source of nucleosomes containing a single CENP-A in our

preparations is that CENP-A complexes may transition from one to two copies of CENP-A in a cell cycle dependent manner.

To test this hypothesis, we performed single-molecule counting on extracts prepared from synchronized cell cultures arrested at various stages of the cell cycle. Cell cycle arrest for each condition was confirmed by flow cytometry (Figure 2C). Similar to asynchronous cells, nucleosomes isolated from CENP-A-YFP cells arrested in G1, early S, late S, G2 and mitosis displayed predominantly two photobleaching events (Figure 2B). Interestingly, compared to asynchronous cells (Figure 1D), the ratio of double bleaching events was slightly lower in all cell cycle arrested samples. Because these samples were prepared and analyzed in the same manner, it is possible that methods for synchronizing cells changed the population bias. Regardless, the overall trend of two CENP-A molecules per complex is still clearly evident. In sum, these data suggest that centromeric nucleosomes contain two molecules of CENP-A throughout the cell cycle and do not support the hypothesis that CENP-A complexes globally transition in a cell cycle dependent manner. An implication from these results is that mechanisms for centromere propagation and kinetochore assembly recognize a similar CENP-A dimer generated structure. Based on our current understanding of these mechanisms, different proteins mediate centromere propagation and kinetochore assembly. It will be of interest to determine if the mechanism of centromere recognition by proteins required for these diverse functions is in fact similar.

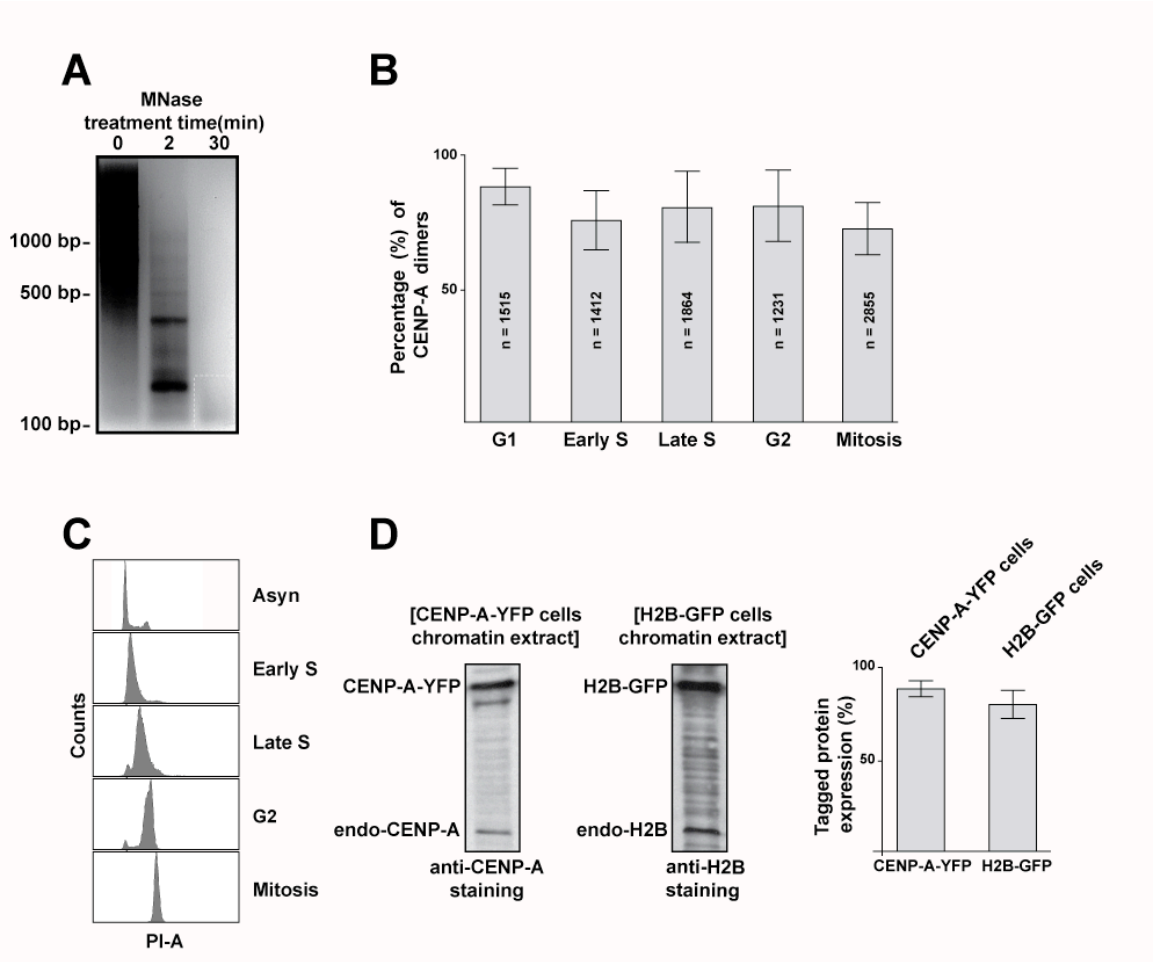

Figure -2 : CENP-A dimers are found at centromeric nucleosomes throughout the cell cycle.

A. Agarose gel of DNA extracted from non-digested and micrococcal nuclease (MNase) digested chromatin lysates (2 min and 30 min treatments) from CENP-A-YFP cells. Note the band around 150 bp, corresponding to the length of DNA wrapped around nucleosomes, protected from micrococcal nuclease digestion (dashed box). **B.** Analysis of nucleosomes derived from various time points in the cell cycle indicated that similar to the asynchronous cells, CENP-A dimers are found at these nucleosomes throughout the cell cycle. **C.** Flow cytometry analysis of propidium iodide staining for DNA content to verify cell cycle arrested

cell populations. **D.** Western blot analysis of the nuclear lysates from CENP-A-YFP and H2B-GFP cells indicated that the tagged proteins are expressed at a higher level compared to the endogenous untagged proteins. (error bars indicate standard deviation).

CENP-A is predominantly complexed in canonical nucleosomes:

Base on the above results, two CENP-A molecules generate the centromere epigenetic mark that regulates chromosome segregation. However, from our results it is not clear that CENP-A is complexed in a nucleosome or in some other structure. This is a key question in understanding how the centromere epigenetic mark is interpreted to ensure genome stability. Indeed, a second difference in models of CENP-A chromatin composition is the presence or absence of other components. Distinct protein constituents could act as an epigenetic mark. Specifically, the chaperone HJURP is thought to replace H2A/H2B in some models representing a substantial structural, possibly epigenetic, change. Based on our results above, the hypothesis of a H4:CENP-A:HJURP trimer is excluded, however a homo-hexamer $(H4:CENP\text{-}A:HJURP)_2$ is not. In such a model, histone H2B or H2A would not be associated with CENP-A.

To determine other components present in CENP-A nucleosomes, we integrated an indirect immunostaining strategy into our PA-CNC assay using isolated nucleosomes. Figure 3A depicts representative immunostaining images. This assay is imperfect due to the non-covalent nature of staining and surface effects on epitope accessibility, as best evidenced by incomplete colocalization between CENP-A-YFP and an antibody directed toward CENP-A (24.3 ± 3.4%) (Figure 3B). With this caveat in mind, co-staining with antibodies to H2B or H4 indicated that a large percentage of individual CENP-A-YFP spots contained H2B and H4 (39.7 ± 3.4% and 42.5 ± 5.9% respectively) (Figure 3B). Thus our single molecule data suggests that centromeric nucleosomes contain all core histones as well as two copies CENP-A and that the structure of these nucleosomes were not compromised during extraction.

CENP-A has also been reported to localize to non-centromeric loci, specifically points of DNA damage in human cells (Zeitlin et al., 2009) and generally along the chromosome arms in situations of CENP-A over-expression (Tomonaga et al., 2003; Van Hooser et al., 2001). Based on many observations over many cellular generations, CENP-A-YFP was never visually detected in non-centromere regions in our cell line, however we cannot exclude that a percentage of non-centromeric CENP-A, originating from chromosome arms, exists in our sample. To determine if the CENP-A analyzed was centromere derived, we used our TIRF based immunostaining assay. CENP-C localizes to centromeres throughout the cell cycle and is required for both centromere and kinetochore regulation (Moree et al., 2011). We reasoned that colocalization with CENP-C would substantiate the idea that the CENP-A population measured are of centromeric origin. Using our TIRF assay, only approximately 17% of CENP-A spots colocalized with CENP-C (compared to less than 5% in H2B controls) (Figure 3B). However, of the CENP-C colocalizing spots, nearly all (91.8 ±7.3%) had two copies of CENP-A indicating that regardless of the percentage of truly centromere derived CENP-A in our samples, analysis of CENP-A number in individual complexes using our assay is consistent. (Figure 3C). Thus, we conclude that centromere derived CENP-A exists as a nucleosome and, as CENP-A is the only core component different from canonical nucleosomes, the mechanisms detecting the centromere epigenetic identity rely on a CENP-A generated mark.

The chaperone HJURP mediates a transition from one to two CENP-A molecules per complex:

Despite our colocalization analysis, our data reproducibly suggest that a minor population of CENP-A particles contain a single copy of CENP-A. In order to further

characterize the secondary small population of spots with a single CENP-A-YFP molecule, we considered two possibilities: first, the biochemical preparation of mono-nucleosomes could somehow have resulted in the bleaching or destruction of individual fluorophores prior to image acquisition, which we term "pre-bleaching". Second, the single CENP-A-YFP spots could be pre-nucleosomal intermediates of CENP-A.

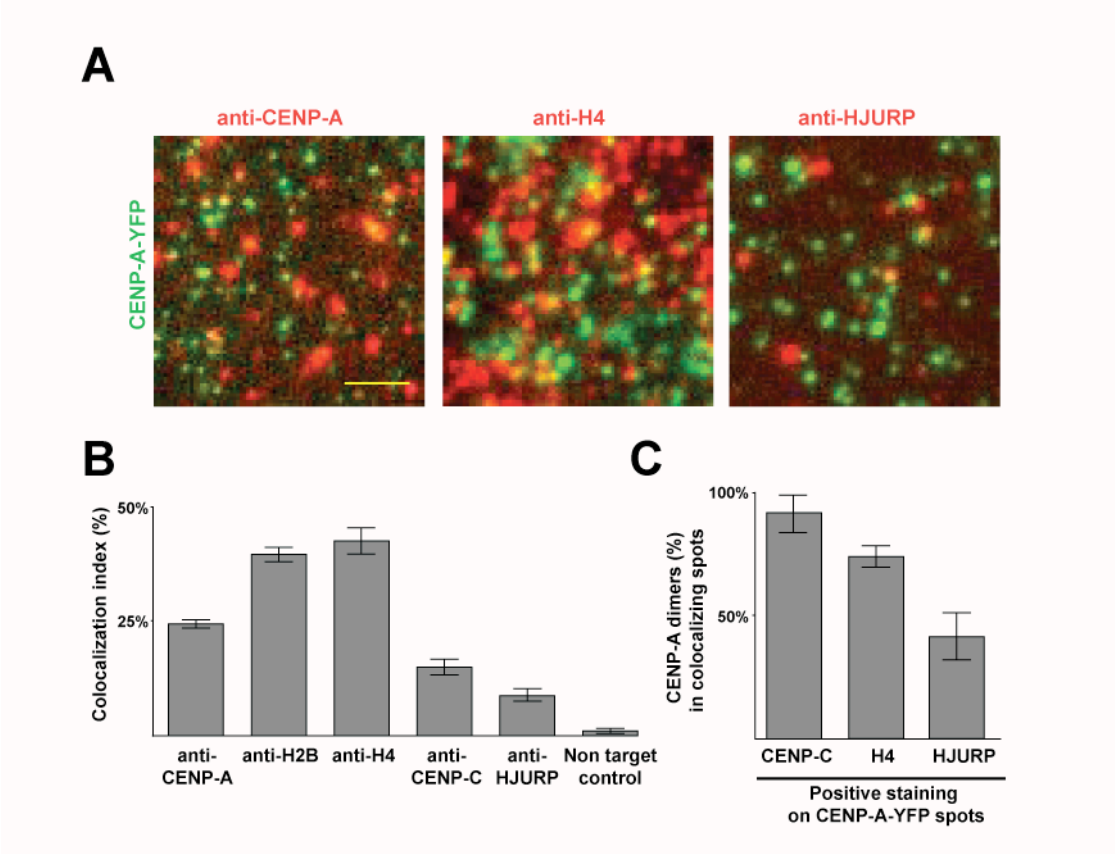

Figure 3 : Core histone H4 and chaperone HJURP colocalize with CENP-A at centromeric nucleosomes and pre-nucleosomal complexes respectively.

A. Representative images of immunofluorescence with anti-H4, anti-HJURP and anti-CENP-A antibodies (red), on CENP-A-YFP nucleosomes (green). **B.** Quantification of colocalization

as number of colocalizing spots divided by number of green spots. **C.** Quantification of bleaching events in CENP-A-YFP spots colocalizing with CENP-C, H4 and HJURP immunofluorescence. (scale bars = 20 μm; error bars indicate standard deviation).

To address the pre-bleaching hypothesis, we imaged GST-GFP dimers in our PA-CNC assay, where we expected nearly exclusively double bleaching events due to the inherent nature of GST to dimerize. We confirmed by pull-down that nearly all the GST-GFP was in dimer form (as detected by western blot, data not shown), as would be expected. However, there was a small but significant population of single bleaching events in the GST-GFP sample (see methods). We considered this small secondary population a result of non-excitation induced pre-bleaching. Therefore we corrected our data for pre-bleached fluorophores based on GST-GFP bleaching events (see methods). However it remains possible that we underestimated pre-bleaching in CENP-A-YFP chromatin lysate due to different isolation methods compared to purified GST-GFP.

To test the second hypothesis, we performed colocalization analysis using an anti-HJURP antibody, as HJURP is the CENP-A:H4 chaperone and expected to bind to preassembled complexes. HJURP staining colocalized with CENP-A-YFP at a low level compared to histone proteins (9.2 ± 2.8%) (Figure 3B). We then performed the PA-CNC technique on colocalizing spots. Interestingly, HJURP-positive CENP-A-YFP spots showed a lower proportion of double-bleaching events compared to CENP-C and H4 positive spots (Figure 3C; 41.9 ± 9.2% versus 91.8± 7.3% and 74.2 ± 4.2%, respectively). These results suggest that while the majority of incorporated CENP-A nucleosomes have an octameric structure, chaperone bound CENP-A may exist in different structures having one or two copies of CENP-A.

Conclusions:

Centromere identity in metazoans is generally accepted to be epigenetically marked by the presence of CENP-A chromatin. Our data presented here are in agreement with the crystal structure of the CENP-A nucleosomes reconstituted *in vitro* from bacterially purified core histones suggesting an octameric nature for centromeric nucleosomes. The centromere epigenetic mark is interpreted by at least two critical mechanisms to ensure genome stability. First, CENP-A chromatin must be detected by proteins required for kinetochore assembly in mitosis. This mechanism is as yet unknown, however culminates in the formation of a single kinetochore on each chromatid that attaches to spindle microtubules, segregating the genome accurately to the daughter cells. Second, during G1, new CENP-A is deposited in centromeres in preparation for DNA replication and equal distribution of the centromere epigenetic mark to the daughter strands. This process, to the best of our knowledge, begins with recognition of CENP-A chromatin by the centromere licensing complex (CLC, made up of KNL-2/M18BP1, Mis18a and Mis18b) (Fujita et al., 2007) by an as yet unknown mechanism. The CLC recruits CENP-A coupled to HJURP (Foltz et al., 2009; Shuaib et al., 2010) for new CENP-A chromatin formation. Importantly, our data suggest that the nature of the centromere epigenetic mark does not grossly change in over the cell cycle. Therefore, these two seemingly independent mechanisms interpret a single species of CENP-A chromatin. It will be of great interest to determine these mechanisms as they likely represent key steps in understanding both chromosome segregation and epigenetic regulation.

Acknowledgments:

We are grateful to Dr. Amy S Maddox, Valérie De Rop, and Rajesh Ranjan for fruitful discussions. We also thank IRIC's medicinal chemistry research unit and Dan Deon for technical assistance. We thank Dr.s Dan Foltz, Alain Verreault, and Iain Cheeseman for the kind gift of anti-HJURP, -H4, and -CENP-C antibodies. P.S.M. is the Canada Research Chair in Cell Division and Chromosomal Organization and supported by research grants from the CIHR (MOP-106548) and CCSRI (700824).

Experimental Procedures:

Cell lines and nuclear lysate preparation

HeLa cells stably expressing CENP-A-YFP or H2B-GFP, as well as plain and transfected HeLa cells were used in this study with standard culturing conditions. For synchronization, cells were collected at different time points following thymidine and nocodazole treatments. DNA of synchronized cell populations was stained with propidium iodide, and cells were analysed by flow cytometry. Transfections were performed using Lipofectamine 2000 reagent (Invitrogen) following manufacturer's instructions.

Crude nuclear lysates were extracted using a Digitonin lysis buffer supplemented with a cocktail of proteinase inhibitors. Briefly, trypsinized cells were incubated with the Digitonin buffer, homogenized by douncing for 30 times and centrifuged at 100g for 3 min to remove cell debris. These steps were repeated 3 times and terminated by centrifugation at 4000g for 15 min to obtain a crude nuclear lysate. The precipitate containing the nuclei was then re-suspended in the initial buffer and kept at -80°C until use. Aliquots of isolated nuclear lysates at a total concentration of 3.6 µg were then treated with 100 units of micrococcal nuclease (Fermentas), in a digestion buffer consisting of 1X PBS and 5 mM $CaCl_2$ and incubated at 37°C for 30 minutes. The reaction was stopped by adding 30 µl of a 0.5 M EDTA solution and chilling of the reaction tube on ice.

To check for complete digestion, DNA was isolated from an aliquot micrococcal nuclease digested nucleosomes. The aliquot was treated with 10 mg/µl RNAse A for 1 hour at 37°C, then 0.5 mg/µl Proteinase K and 0.5 % SDS for 1 hour at 50°C. DNA was then precipitated and eluted on a column (Omega Bio-Tek), and run on a 1% agarose gel.

Cover slips and Flow chamber

Standard microscopy glass cover slips (thickness No. 1.5) were treated with a mixture of sulfuric acid and hydrogen peroxide (Piranha solution) and stored in ultrapure water prior to usage. To make an imaging flow chamber, two pieces of double-sided tape were applied to a glass slide, parallel to each other, over which a cover slip was placed.

10 µl of a 1 mg/ml solution of YFP nanobodies was perfused, three times into the flow chamber, and incubated for 10 minutes, with a standard pipette and followed by 10 µl of 1X PBS to wash. Similarly, 10 µl of digested nuclei was then perfused three times, incubated for 10 minutes and eventually washed with 1X PBS. Slides were imaged immediately after preparation.

Immuno-staining within the flow chamber

Once samples were perfused into flow chambers, primary anti- CENP-A, H2B and cMyc (Abcam at 1mg/ml diluted 1:1000) were perfused, incubated for 10 minutes and washed out with 1X PBS. The HJURP and H4 and CENP-C antibodies, used at the same dilutions, were kind gifts from the laboratories of Drs Daniel Foltz, Alain Verreault and Iain M. Cheeseman. Secondary antibodies (Invitrogen at 1mg/ml diluted 1:1000) conjugated to Cy3 were added to the flow chamber, incubated for 1 min and washed with 30 ul of 1X PBS.

Single-Molecule Microscopy

All imaging was carried out on a Nikon TI Eclipse inverted microscope, equipped with a 100X 1.49 NA APO-TIRF objective (Nikon, Melville, NY), illuminated with a 488, 568, and 647nm laser launch (Nikon). Image acquisition was performed with a Cascade II EMCCD camera (Photometrics), controlled with NIS-Elements software (Nikon). The focus was found

using the 568nm laser to prevent pre-bleaching of fluorescent proteins, and images were acquired using full 488nm laser power (10mW) for 900ms exposures at 1 second intervals for 2-4 minutes. The EMCCD was operated in full frame (unbinned) normal (non-EM) readout mode at the slowest speed to reduce noise.

Image and data analysis

Manual image analysis was performed using ImageJ (NIH), by selecting a subset of spots as ROIs, and extracting average area pixel intensity through time. Intensity profiles over time were then subjected to blind classification based on the number of intensity drops corresponding to bleaching events. Automated quantification of bleaching events was performed using a custom analysis software written in Matlab (Mathworks), involving a pipeline of multiple algorithms. To find the spots, images were segmented using a probabilistic approach. Then, spots were tracked using the u-track framework. Finally, bleaching events were detected by first identifying putative drops by peak detection in the first derivative of median-filtered intensity profiles, followed by applying a t-test to determine whether intensities before and after putative drops were different with a 95% confidence interval. Colocalization analysis was performed in Matlab by applying segmentation to images, and counting overlapping spots.

Quantification of relative protein expression levels of unlabeled endogenous protein vs. labelled protein was performed by integrating pixel density of corresponding bands on western blot images, using ImageJ (NIH).

Correction of apparent counts

Our goal is to determine the fraction *R* of complexes containing two copies of CENP-A given the number of complexes with two copies of CENP-A (Δ) and the number of complexes with one copy of CENP-A (Σ):

$$R = \frac{\Delta}{\Delta + \Sigma}$$

However, neither Δ nor Σ are directly observable, due to both incomplete labeling of molecules, and photobleaching events before the start of image acquisition. Consequently, we need to apply a correction to the observed numbers of profiles with two bleach events ("doubles"; *D*) and with one bleach event ("singles"; *S*), respectively, in order to obtain the correct value for *R*.

With a labeled fraction of CENP-A molecules f_L (90% in our cell line), we can calculate the number of doubles and singles we expect as

$$D_0 = f_L^2 \Delta$$

$$S_0 = 2 f_L (1 - f_L) \Delta + f_L \Sigma .$$

In addition to incomplete labeling, the amount of fluorescent CENP-A is further reduced due to the destruction of fluorophores before imaging starts, which can occur throughout the process of sample preparation. With the fraction of "pre-bleached" molecules f_B (f_B of 0.2 would mean that 20% of previously fluorescent molecules have bleached – in our hands, f_B is typically 12%), we find the expected numbers of singles and doubles as

$$D_1 = (1 - f_B)^2 f_L^2 \Delta$$

$$S_1 = 2f_B(1 - f_B)f_L^2\Delta + (1 - f_B)(2f_L(1 - f_L)\Delta + f_L\Sigma).$$

Thus, we can express the fraction of complexes containing 2 copies of CENP-A as a function of the observed numbers of doubles and singles given incomplete labeling and pre-bleaching as

$$R = \frac{D_1}{(2(1 - f_B)f_L - 1)D_1 + (1 - f_B)f_L S_1}.$$

To estimate f_B from GST-GFP controls, where $R=1$ and $L=1$, we use

$$f_B = \frac{S_1}{S_1 + 2D_1}.$$

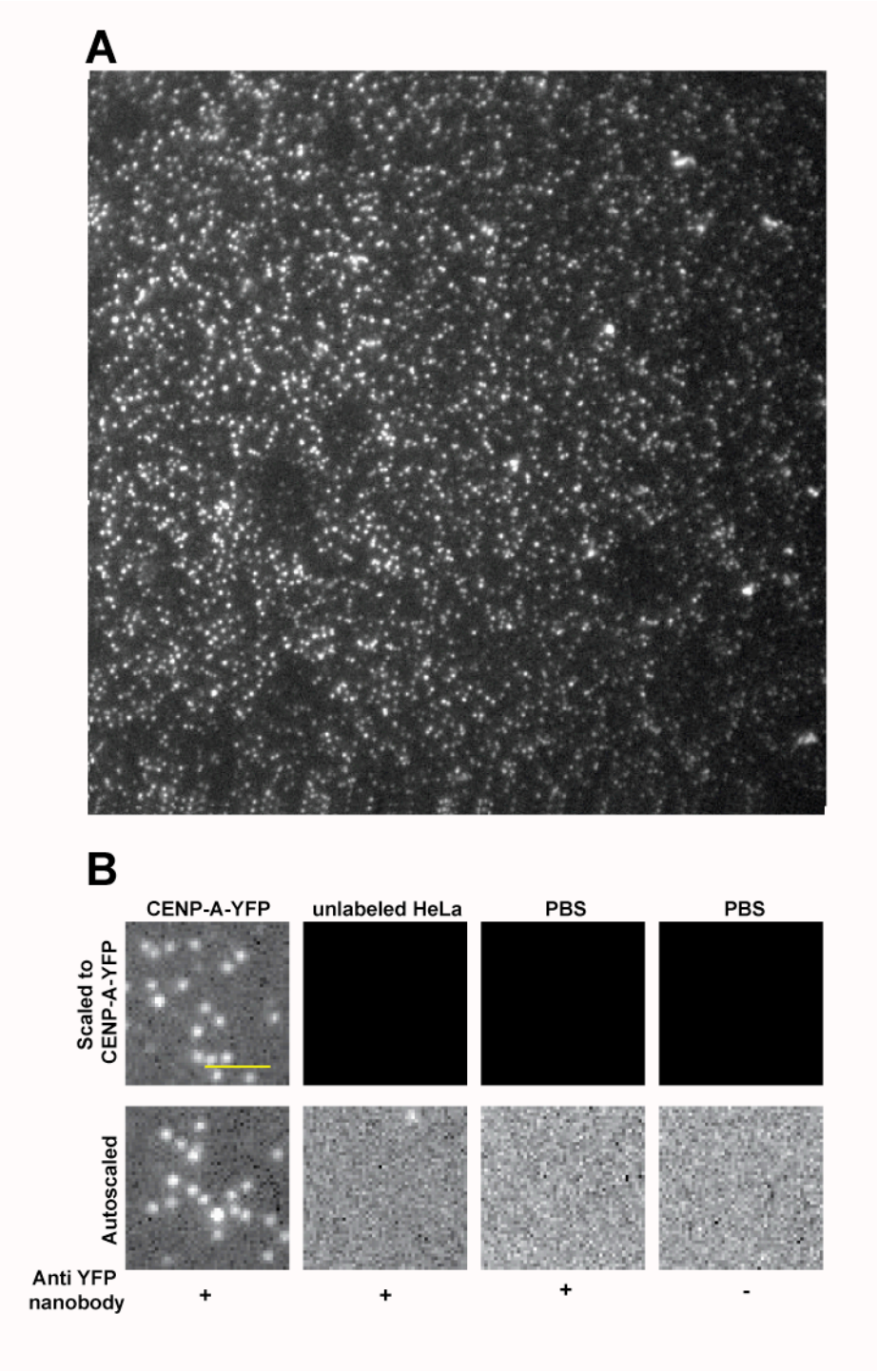

Figure 4 : Supplement I

A. A Typical TIRFM image of pulled-down CENP-A-YFP nucleosomes, full 512x512 pixel frame. **B.** Zoomed in 50x50 pixel images of CENP-A-YFP nucleosomes, as well as nuclear lysate from untagged HeLa cells, and PBS control perfusions with and without GBP. Top row images are scaled to CENP-A-YFP to show contrast between images, and bottom row images are autoscaled to show contrast within images. (scale bars = 20 μm;)

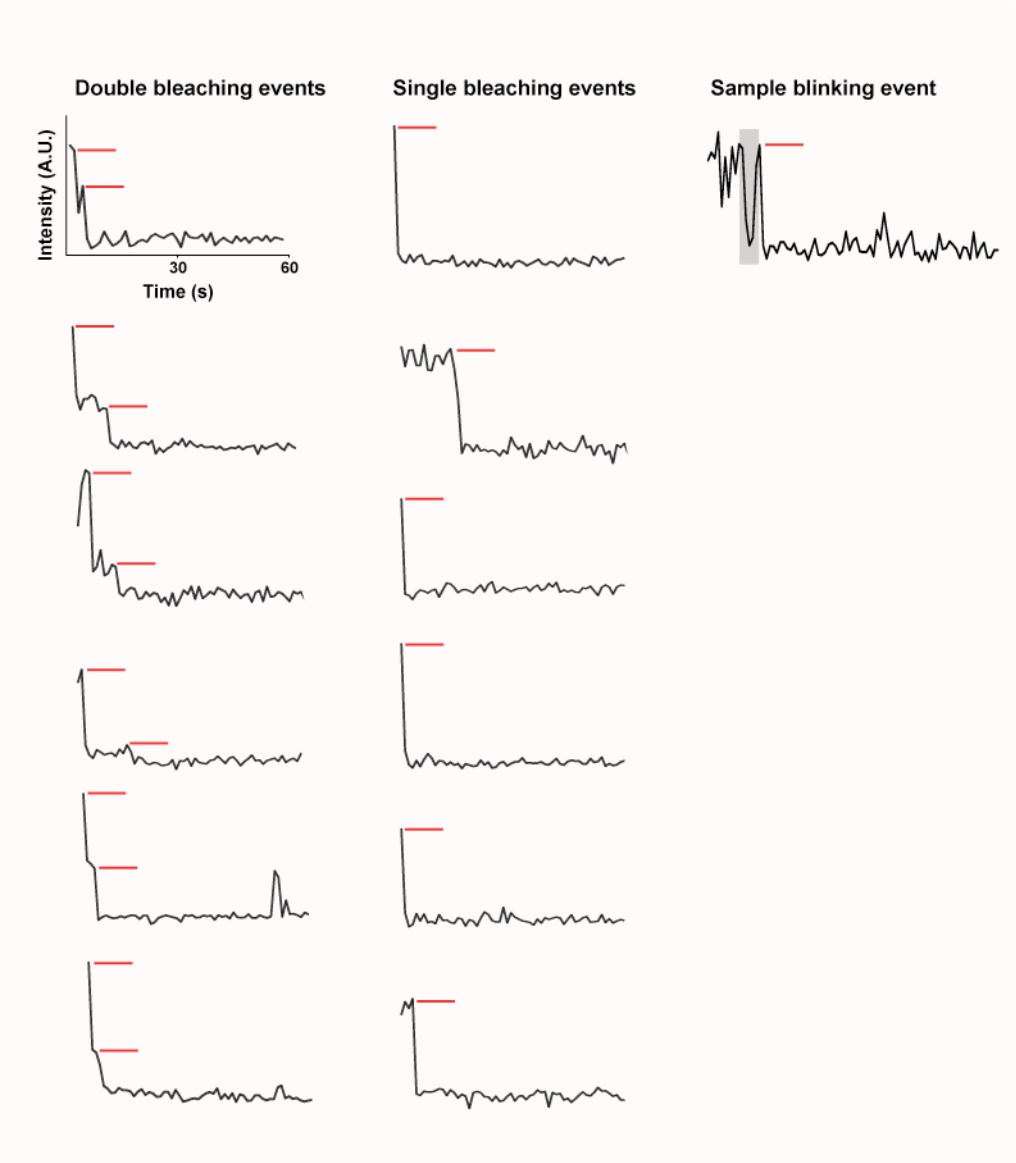

Figure 5 : Supplement II

Example traces of intensity fluctuations over time, classified as displaying two or one bleaching events as well as a sample blinking event. Red lines indicate plateaus of intensity, which allow visualization of bleaching steps. Right, example trace of blinking event (highlighted in grey), where intensity briefly falls to baseline intensity, and recovers to initial plateau.

References:

Allshire, R.C., and Karpen, G.H. (2008). Epigenetic regulation of centromeric chromatin: old dogs, new tricks? Nat Rev Genet *9*, 923-937.

Barnhart, M.C., Kuich, P.H., Stellfox, M.E., Ward, J.A., Bassett, E.A., Black, B.E., and Foltz, D.R. (2011). HJURP is a CENP-A chromatin assembly factor sufficient to form a functional de novo kinetochore. J Cell Biol *194*, 229-243.

Bassett, E.A., DeNizio, J., Barnhart-Dailey, M.C., Panchenko, T., Sekulic, N., Rogers, D.J., Foltz, D.R., and Black, B.E. (2012). HJURP uses distinct CENP-A surfaces to recognize and to stabilize CENP-A/histone H4 for centromere assembly. Dev Cell *22*, 749-762.

Bernad, R., Sanchez, P., Rivera, T., Rodriguez-Corsino, M., Boyarchuk, E., Vassias, I., Ray-Gallet, D., Arnaoutov, A., Dasso, M., Almouzni, G.*, et al.* (2011). Xenopus HJURP and condensin II are required for CENP-A assembly. The Journal of cell biology *192*, 569-582.

Black, B.E., Brock, M.A., Bedard, S., Woods, V.L., Jr., and Cleveland, D.W. (2007a). An epigenetic mark generated by the incorporation of CENP-A into centromeric nucleosomes. Proc Natl Acad Sci U S A *104*, 5008-5013.

Black, B.E., Foltz, D.R., Chakravarthy, S., Luger, K., Woods, V.L., Jr., and Cleveland, D.W. (2004). Structural determinants for generating centromeric chromatin. Nature *430*, 578-582.

Black, B.E., Jansen, L.E., Maddox, P.S., Foltz, D.R., Desai, A.B., Shah, J.V., and Cleveland, D.W. (2007b). Centromere identity maintained by nucleosomes assembled with histone H3 containing the CENP-A targeting domain. Mol Cell *25*, 309-322.

Bui, M., Dimitriadis, E.K., Hoischen, C., An, E., Quenet, D., Giebe, S., Nita-Lazar, A., Diekmann, S., and Dalal, Y. (2012). Cell-Cycle-Dependent Structural Transitions in the Human CENP-A Nucleosome In Vivo. Cell *150*, 317-326.

Bulseco, D.A., and Wolf, D.E. (2007). Fluorescence correlation spectroscopy: molecular complexing in solution and in living cells. Methods in cell biology *81*, 525-559.

Camahort, R., Shivaraju, M., Mattingly, M., Li, B., Nakanishi, S., Zhu, D., Shilatifard, A., Workman, J.L., and Gerton, J.L. (2009). Cse4 is part of an octameric nucleosome in budding yeast. Mol Cell *35*, 794-805.

Cho, U.S., and Harrison, S.C. (2011). Recognition of the centromere-specific histone Cse4 by the chaperone Scm3. Proc Natl Acad Sci U S A *108*, 9367-9371.

Cleveland, D.W., Mao, Y., and Sullivan, K.F. (2003). Centromeres and kinetochores: from epigenetics to mitotic checkpoint signaling. Cell *112*, 407-421.

Coffman, V.C., Wu, P., Parthun, M.R., and Wu, J.Q. (2011). CENP-A exceeds microtubule attachment sites in centromere clusters of both budding and fission yeast. J Cell Biol *195*, 563-572.

Conde e Silva, N., Black, B.E., Sivolob, A., Filipski, J., Cleveland, D.W., and Prunell, A. (2007). CENP-A-containing nucleosomes: easier disassembly versus exclusive centromeric localization. J Mol Biol *370*, 555-573.

Dalal, Y., Wang, H., Lindsay, S., and Henikoff, S. (2007). Tetrameric structure of centromeric nucleosomes in interphase Drosophila cells. PLoS Biol *5*, e218.

De Rop, V., Padeganeh, A., and Maddox, P.S. (2012). CENP-A: the key player behind centromere identity, propagation, and kinetochore assembly. Chromosoma.

Dimitriadis, E.K., Weber, C., Gill, R.K., Diekmann, S., and Dalal, Y. (2010). Tetrameric organization of vertebrate centromeric nucleosomes. Proc Natl Acad Sci U S A *107*, 20317-20322.

Dunleavy, E.M., Roche, D., Tagami, H., Lacoste, N., Ray-Gallet, D., Nakamura, Y., Daigo, Y., Nakatani, Y., and Almouzni-Pettinotti, G. (2009). HJURP is a cell-cycle-dependent maintenance and deposition factor of CENP-A at centromeres. Cell *137*, 485-497.

Foltz, D.R., Jansen, L.E., Bailey, A.O., Yates, J.R., 3rd, Bassett, E.A., Wood, S., Black, B.E., and Cleveland, D.W. (2009). Centromere-specific assembly of CENP-a nucleosomes is mediated by HJURP. Cell *137*, 472-484.

Fujita, Y., Hayashi, T., Kiyomitsu, T., Toyoda, Y., Kokubu, A., Obuse, C., and Yanagida, M. (2007). Priming of centromere for CENP-A recruitment by human hMis18alpha, hMis18beta, and M18BP1. Dev Cell *12*, 17-30.

Fukagawa, T. (2004). Centromere DNA, proteins and kinetochore assembly in vertebrate cells. Chromosome Res *12*, 557-567.

Furuyama, T., and Henikoff, S. (2009). Centromeric nucleosomes induce positive DNA supercoils. Cell *138*, 104-113.

Guse, A., Carroll, C.W., Moree, B., Fuller, C.J., and Straight, A.F. (2011). In vitro centromere and kinetochore assembly on defined chromatin templates. Nature *477*, 354-358.

Hayashi, T., Fujita, Y., Iwasaki, O., Adachi, Y., Takahashi, K., and Yanagida, M. (2004). Mis16 and Mis18 are required for CENP-A loading and histone deacetylation at centromeres. Cell *118*, 715-729.

Henikoff, S., and Henikoff, J.G. (2012). "Point" centromeres of Saccharomyces harbor single centromere-specific nucleosomes. Genetics *190*, 1575-1577.

Howman, E.V., Fowler, K.J., Newson, A.J., Redward, S., MacDonald, A.C., Kalitsis, P., and Choo, K.H. (2000). Early disruption of centromeric chromatin organization in centromere protein A (Cenpa) null mice. Proc Natl Acad Sci U S A *97*, 1148-1153.

Hu, H., Liu, Y., Wang, M., Fang, J., Huang, H., Yang, N., Li, Y., Wang, J., Yao, X., Shi, Y.*, et al.* (2011). Structure of a CENP-A-histone H4 heterodimer in complex with chaperone HJURP. Genes Dev *25*, 901-906.

Huang, C.C., Chang, K.M., Cui, H., and Jayaram, M. (2011). Histone H3-variant Cse4-induced positive DNA supercoiling in the yeast plasmid has implications for a plasmid origin of a chromosome centromere. Proc Natl Acad Sci U S A *108*, 13671-13676.

Jain, A., Liu, R., Ramani, B., Arauz, E., Ishitsuka, Y., Ragunathan, K., Park, J., Chen, J., Xiang, Y.K., and Ha, T. (2011). Probing cellular protein complexes using single-molecule pull-down. Nature *473*, 484-488.

Jansen, L.E., Black, B.E., Foltz, D.R., and Cleveland, D.W. (2007). Propagation of centromeric chromatin requires exit from mitosis. The Journal of cell biology *176*, 795-805.

Krassovsky, K., Henikoff, J.G., and Henikoff, S. (2012). Tripartite organization of centromeric chromatin in budding yeast. Proc Natl Acad Sci U S A *109*, 243-248.

Krichevsky, O., and Bonnet, G. (2002). Fluorescence correlation spectroscopy: the technique and its applications. Rep Prog Phys *65*, 251-297.

Lagana, A., Dorn, J.F., De Rop, V., Ladouceur, A.M., Maddox, A.S., and Maddox, P.S. (2010). A small GTPase molecular switch regulates epigenetic centromere maintenance by stabilizing newly incorporated CENP-A. Nat Cell Biol *12*, 1186-1193.

Lavelle, C., and Prunell, A. (2007). Chromatin polymorphism and the nucleosome superfamily: a genealogy. Cell Cycle *6*, 2113-2119.

Lawrimore, J., Bloom, K.S., and Salmon, E.D. (2011). Point centromeres contain more than a single centromere-specific Cse4 (CENP-A) nucleosome. J Cell Biol *195*, 573-582.

Lochmann, B., and Ivanov, D. (2012). Histone H3 localizes to the centromeric DNA in budding yeast. PLoS Genet *8*, e1002739.

Maddox, P.S., Hyndman, F., Monen, J., Oegema, K., and Desai, A. (2007). Functional genomics identifies a Myb domain-containing protein family required for assembly of CENP-A chromatin. The Journal of cell biology *176*, 757-763.

Mellone, B.G., Grive, K.J., Shteyn, V., Bowers, S.R., Oderberg, I., and Karpen, G.H. (2011). Assembly of Drosophila centromeric chromatin proteins during mitosis. PLoS Genet *7*, e1002068.

Mizuguchi, G., Xiao, H., Wisniewski, J., Smith, M.M., and Wu, C. (2007). Nonhistone Scm3 and histones CenH3-H4 assemble the core of centromere-specific nucleosomes. Cell *129*, 1153-1164.

Moree, B., Meyer, C.B., Fuller, C.J., and Straight, A.F. (2011). CENP-C recruits M18BP1 to centromeres to promote CENP-A chromatin assembly. The Journal of cell biology *194*, 855-871.

Nishino, T., Takeuchi, K., Gascoigne, K.E., Suzuki, A., Hori, T., Oyama, T., Morikawa, K., Cheeseman, I.M., and Fukagawa, T. (2012). CENP-T-W-S-X forms a unique centromeric chromatin structure with a histone-like fold. Cell *148*, 487-501.

Oudet, P., Germond, J.E., Bellard, M., Spadafora, C., and Chambon, P. (1978). Nucleosome structure. Philos Trans R Soc Lond B Biol Sci *283*, 241-258.

Padeganeh, A., De Rop, V., and Maddox, P.S. (2013). Nucleosomal composition at the centromere: a numbers game. Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology.

Palmer, D.K., O'Day, K., Trong, H.L., Charbonneau, H., and Margolis, R.L. (1991). Purification of the centromere-specific protein CENP-A and demonstration that it is a distinctive histone. Proc Natl Acad Sci U S A *88*, 3734-3738.

Panchenko, T., Sorensen, T.C., Woodcock, C.L., Kan, Z.Y., Wood, S., Resch, M.G., Luger, K., Englander, S.W., Hansen, J.C., and Black, B.E. (2011). Replacement of histone H3 with CENP-A directs global nucleosome array condensation and loosening of nucleosome superhelical termini. Proc Natl Acad Sci U S A *108*, 16588-16593.

Pearson, C.G., Maddox, P.S., Salmon, E.D., and Bloom, K. (2001). Budding yeast chromosome structure and dynamics during mitosis. The Journal of cell biology *152*, 1255-1266.

Phillips, G.N., Jr. (1997). Structure and dynamics of green fluorescent protein. Curr Opin Struct Biol *7*, 821-827.

Probst, A.V., Dunleavy, E., and Almouzni, G. (2009). Epigenetic inheritance during the cell cycle. Nat Rev Mol Cell Biol *10*, 192-206.

Reeve, J.N., Sandman, K., and Daniels, C.J. (1997). Archaeal histones, nucleosomes, and transcription initiation. Cell *89*, 999-1002.

Rothbauer, U., Zolghadr, K., Muyldermans, S., Schepers, A., Cardoso, M.C., and Leonhardt, H. (2008). A versatile nanotrap for biochemical and functional studies with fluorescent fusion proteins. Mol Cell Proteomics *7*, 282-289.

Rust, M.J., Bates, M., and Zhuang, X. (2006). Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). Nature methods *3*, 793-795.

Schuh, M., Lehner, C.F., and Heidmann, S. (2007). Incorporation of Drosophila CID/CENP-A and CENP-C into centromeres during early embryonic anaphase. Curr Biol *17*, 237-243.

Sekulic, N., Bassett, E.A., Rogers, D.J., and Black, B.E. (2010). The structure of (CENP-A-H4)(2) reveals physical features that mark centromeres. Nature *467*, 347-351.

Shelby, R.D., Monier, K., and Sullivan, K.F. (2000). Chromatin assembly at kinetochores is uncoupled from DNA replication. J Cell Biol *151*, 1113-1118.

Shivaraju, M., Unruh, J.R., Slaughter, B.D., Mattingly, M., Berman, J., and Gerton, J.L. (2012). Cell-cycle-coupled structural oscillation of centromeric nucleosomes in yeast. Cell *150*, 304-316.

Shuaib, M., Ouararhni, K., Dimitrov, S., and Hamiche, A. (2010). HJURP binds CENP-A via a highly conserved N-terminal domain and mediates its deposition at centromeres. Proc Natl Acad Sci U S A *107*, 1349-1354.

Sullivan, K.F., Hechenberger, M., and Masri, K. (1994). Human CENP-A contains a histone H3 related histone fold domain that is required for targeting to the centromere. J Cell Biol *127*, 581-592.

Tachiwana, H., Kagawa, W., Shiga, T., Osakabe, A., Miya, Y., Saito, K., Hayashi-Takanaka, Y., Oda, T., Sato, M., Park, S.Y.*, et al.* (2011). Crystal structure of the human centromeric nucleosome containing CENP-A. Nature *476*, 232-235.

Tachiwana, H., and Kurumizaka, H. (2011). Structure of the CENP-A nucleosome and its implications for centromeric chromatin architecture. Genes Genet Syst *86*, 357-364.

Tomonaga, T., Matsushita, K., Yamaguchi, S., Oohashi, T., Shimada, H., Ochiai, T., Yoda, K., and Nomura, F. (2003). Overexpression and mistargeting of centromere protein-A in human primary colorectal cancer. Cancer Res *63*, 3511-3516.

Van Hooser, A.A., Ouspenski, II, Gregson, H.C., Starr, D.A., Yen, T.J., Goldberg, M.L., Yokomori, K., Earnshaw, W.C., Sullivan, K.F., and Brinkley, B.R. (2001). Specification of kinetochore-forming chromatin by the histone H3 variant CENP-A. J Cell Sci *114*, 3529-3542.

Weintraub, H., Worcel, A., and Alberts, B. (1976). A model for chromatin based upon two symmetrically paired half-nucleosomes. Cell *9*, 409-417.

Yoda, K., Ando, S., Morishita, S., Houmura, K., Hashimoto, K., Takeyasu, K., and Okazaki, T. (2000). Human centromere protein A (CENP-A) can replace histone H3 in nucleosome reconstitution in vitro. Proc Natl Acad Sci U S A *97*, 7266-7271.

Zeitlin, S.G., Baker, N.M., Chapados, B.R., Soutoglou, E., Wang, J.Y., Berns, M.W., and Cleveland, D.W. (2009). Double-strand DNA breaks recruit the centromeric histone

CENP-A. Proceedings of the National Academy of Sciences of the United States of America *106*, 15762-15767.

Zhang, W., Colmenares, S.U., and Karpen, G.H. (2012). Assembly of Drosophila centromeric nucleosomes requires CID dimerization. Mol Cell *45*, 263-269.

Zhou, Z., Feng, H., Zhou, B.R., Ghirlando, R., Hu, K., Zwolak, A., Miller Jenkins, L.M., Xiao, H., Tjandra, N., Wu, C.*, et al.* (2011). Structural basis for recognition of centromere histone variant CenH3 by the chaperone Scm3. Nature *472*, 234-237.

Chapter 3

Centromere changes upon terminal differentiation in mouse cells

Abbas Padeganeh[1], Joel Ryan[1] and Paul S Maddox[1,2]

[1]Institute for Research in Immunology and Cancer (IRIC), Faculty of Medicine, University of Montreal, H3T 1J4, Quebec, Canada

[2]Department of Biology, University of North Carolina at Chapel Hill, Coker Hall, Campus Box 3280, Chapel Hill, NC 27599-3280, USA

In this chapter, I describe the cellular fate of the centromere epigenetic mark in terminal differentiation when cells exit the cell cycle.

The material presented herein is a manuscript to be submitted to ***Cell cycle.***

Author contribution:

PSM and AP designed the project. AP performed the wet lab experiments, wrote the manuscript and prepared the images. JR assisted with data analysis and PSM edited the manuscript.

Abstract:

The histone H3 variant, CENP-A, is a key component in the epigenetic specification of centromere identity in most species. In proliferating cells, CENP-A chromatin is particularly considered active during centromere inheritance in S phase, CENP-A reloading in G1 and the proteinaceous kinetochore assembly in M phase. Given its important role in mediating chromosome segregation thereby maintaining ploidy, the structure, dynamics and regulation of CENP-A chromatin during the cell cycle have been extensively studied. However, the fate of CENP-A chromatin in non-cycling differentiated cells remains elusive. Here, we report the differential expression and nuclear localization of CENP-A in various adult mouse tissue sections. Interestingly, we also demonstrate that during myogenic differentiation, centromeres undergo global changes including clustering, extensive loss of CENP-A, spatial rearrangement and association of the clusters to large heterochromatin regions. We propose that these processes ensure the conservation of a minimal epigenetic mark, in an inactive state in post-mitotic cells.

Introduction:

The presence of an intact centromere on each chromosome is a key factor in the assembly of the proteinaceous kinetochore thereby faithful partitioning of the genetic material during each round of cell division (Cleveland et al., 2003). Centromere protein-A (CENP-A), a centromere-specific histone H3 variant, is prominently recognized as the epigenetic identifier of centromeric entity in most organisms (De Rop et al., 2012). In the mammalian system, centromeric chromatin is composed of blocks of CENP-A-containing nucleosomes as well as canonical H3 nucleosomes (Dunleavy et al., 2011) together flanked by pericentric heterochromatin (Sullivan and Karpen, 2004), which physically restricts the CENP-A-containing chromatin (Allshire and Karpen, 2008). We have previously demonstrated that in proliferating human cells, octameric CENP-A nucleosomes reside at centromeres in a cell cycle-independent manner (Dunleavy et al., 2013; Padeganeh et al., 2013a; Padeganeh et al., 2013b). In cycling cells, centromeric chromatin is deemed active in certain time points including centromere inheritance is S phase, CENP-A deposition in G1 and kinetochore assembly in M phase (De Rop et al., 2012). The existence of tight regulatory mechanisms for cellular processes involving CENP-A chromatin is crucial and defects in them could potentially compromise ploidy or the normal cell division cycle (Heun et al., 2006; Tomonaga et al., 2003; Valdivia et al., 2009). However, in spite of extensive efforts to dissect the structure, function, dynamics and regulation of CENP-A chromatin in proliferating cells (Black et al., 2007a; Black et al., 2004; Bui et al., 2012; Camahort et al., 2009; Chan and Wong, 2012; Dalal et al., 2007; Dimitriadis et al., 2010; Dunleavy et al., 2009; Krassovsky et al., 2012; Lagana et al., 2010; Mizuguchi et al., 2007; Padeganeh et al., 2013b; Sekulic et al., 2010; Tachiwana et al., 2011; Tomonaga et al., 2003; Van Hooser et al., 2001; Yoda et al.,

2000; Zhou et al., 2011), the fate of CENP-A chromatin outside the cell cycle, e.g. when a tissue-resident stem cell terminally differentiates and permanently exits the cell cycle, remains elusive and largely unexplored. This is particularly important as in a number of human malignancies e.g. colon, breast and lung cancers, CENP-A has been shown to be deregulated where higher CENP-A mRNA and/or protein levels are associated with poor prognosis and higher tumor grades (Amato et al., 2009; McGovern et al., 2012; Tomonaga et al., 2003). Regardless of the underlying mechanism(s), cancer development is universally known to be downstream of uncontrolled cell division, central to which is centromere-mediated chromosome segregation. Therefore, it is presumable that when cells evade cell cycle exit and remain in a proliferative state, CENP-A-containing chromatin is somehow perturbed in which it remains active allowing uncontrolled rounds of proliferation. Given the cell-cycle-specific roles attributed to CENP-A, it has been proposed as an ideal anti-cancer target for the development of small molecule inhibitors (Li et al., 2011; Valdivia et al., 2009). To better understand the fate of CENP-A chromatin in post-mitotic cells, we visualized the centromeric chromatin in the context of various tissue sections and during C2C12 myogenic differentiation. We observed that nuclear localization of centromeres as well as CENP-A levels differ in various tissue contexts. These parameters were also observed to undergo dramatic changes upon C2C12 myogenic differentiation suggesting possible mechanisms for inactivation and conservation of the CENP-A epigenetic mark in terminally differentiated cells.

Results:

CENP-A is expressed differentially in adult tissue sections.

In order to investigate the cellular fate of CENP-A chromatin upon cell cycle exit, we anticipated observing differential phenotypes depending on the rate of cell proliferation depending the tissue origin. Therefore, we performed immunohistochemistry (IHC) on paraffin-embedded adult mouse tissue sections to visualize the centromere epigenetic marker, CENP-A (Figure 1A) and performed quantitative confocal microscopy analysis on CENP-A protein expression and distribution pattern. Interestingly, we observed that in most tissue sections studied, the average number of centromeric foci is much lower than the expected number of diploid chromosomes (n=40), with the average numbers of 2.5 ± 1.8 for heart, 25.3 ± 11 for skin, 6± 1 for brain, 8.8± 2.5 for eye, 2.7 ± 1.3 for uterus and 9.5 ±2.9 for spleen (Figure 1B). Quantitative assessment of signal intensity also revealed significant differences in total CENP-A intensities per nucleus when normalized to the maximum intensity observed (assuming 100% for the highest value) ranging from 2.6 % for heart, 100% for skin, 14.1 % for brain, 83.6 % for eye, 1.2% for uterus and 18.6 % for spleen (Figure 1C). Intriguingly, skin cells are observed to contain the highest number of centromeric foci as well as the highest total CENP-A intensities among the tissue sections under study. Moreover, we also observed that in tissues where the number of centromeric foci was very low e.g. heart and uterus, all centromeric signals were spatially restricted to large DAPI-dense foci that represent heterochromatin regions and thus are known to be enriched for HP1 and H3K9me2/3(Lange et al., 2013; Schotta et al., 2004). These observations in turn suggest that e.g. in case of heart myocytes, each ~16 centromeres cluster together to form ~2.5 visually detectable CENP-A

foci accumulated at large heterochromatin regions collectively hereafter referred to as **HE**terochromatin-localized **C**entromere **C**lusters (HECCs).

Centromeres undergo changes in number, CENP-A levels and distribution pattern upon C2C12 myogenic differentiation:

In order to further investigate the dynamics of CENP-A chromatin upon cell cycle exit, we leveraged the powerful C2C12 myogenic progenitor system (McMahon et al., 1994) (Lawson and Purslow, 2000). This stem cell model allows us to visually track the differentiation process. Individual myoblasts exit the cell cycle and the resultant myocytes fuse to form multinucleated myotubes where the nuclei align together while maintaining their individual entity.

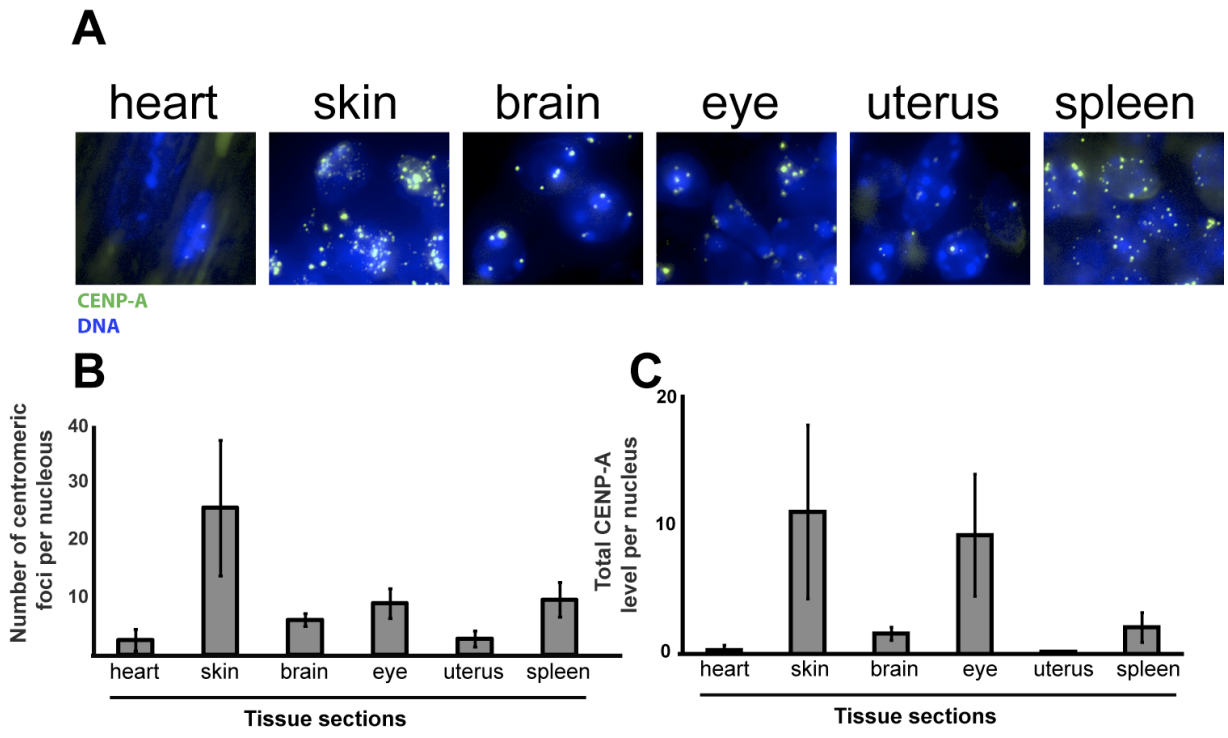

Figure 1 : CENP-A is differentially expressed in adult mouse tissues.

A. Paraffin-embedded sections of various tissue origins from adult mice stained for CENP-A. Note that for the purpose of visual inspection of nuclear distribution and number of CENP-A foci, maximum projection and scaling of the intensities were applied. **B.** Average number of CENP-A signals per nucleus showed dramatic difference between skin and other tissue cells. **C.** Total fluorescence intensity of CENP-A signals per nucleus as quantified from raw confocal microscopy images.

To confirm our differentiation protocol, BrdU incorporation experiments were conducted besides visual inspection of cellular morphology, and as expected, we observed a large number of BrdU positive nuclei in myoblasts whereas in day 3 post-differentiation myotubes there was no detectable BrdU staining indicating that the post-differentiation nuclei had also exited the cell cycle (supplement I).

Interestingly, using immunofluorescence staining (IF), we observed that upon terminal myogenic differentiation of C2C12 myoblasts (Figure 2), the number of centromeric foci per nucleus was significantly reduced in day 3 post-differentiation myotubes with an average number of 65.6 centromeric foci in myoblasts to 24.2 in myotubes (it is noteworthy that this cell line is near tetraploid) (Figure 3B). This phenomenon was also observed in longer incubations of myotubes in differentiation media with average numbers of 31.2, 22.2, 23 and 21 centromeric foci detected in days 6, 9, 12 and 18 post-differentiation (Figure 3B). Interestingly, we also observed a sharp decrease in total CENP-A intensity levels per nucleus in post-mitotic myotubes dropping from 100% in myoblasts to 13.2 %, 22.7 %, 24 %, 19.4 % and 21 % in myotubes of days 3, 6, 9, 12 and 18 post-differentiation respectively (Figure 3C).

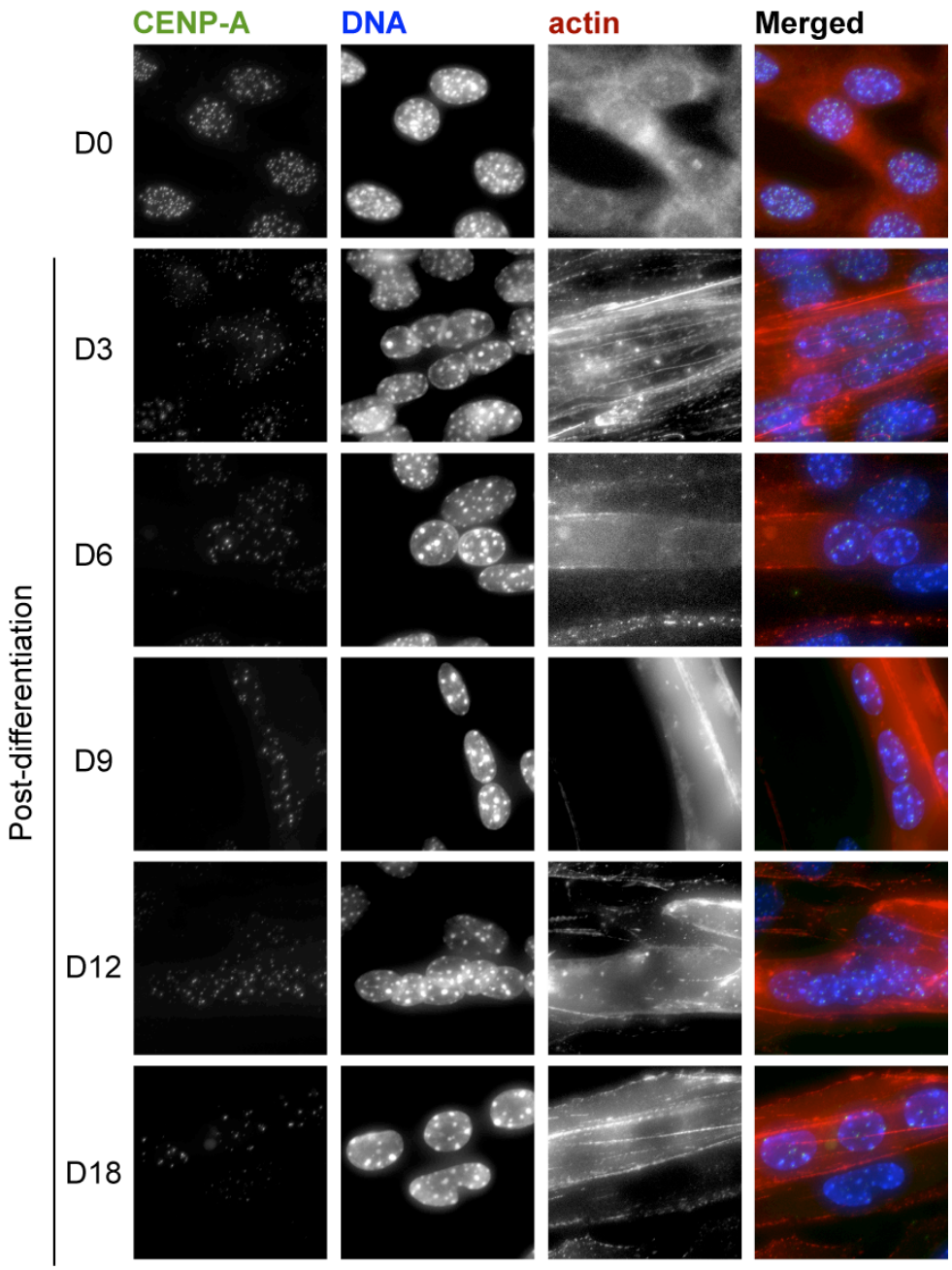

Figure 2 : C2C12 myogenic system as a powerful stem cell model.

C2C12 myoblasts were maintained in the proliferative state by growing them in DMEM media supplemented with 10% FBS prior to the induction of differentiation (Day 0, D0). Differentiation was induced by replacing the media with 1% horse-serum containing DMEM

and images were acquired from fixed myotubes 3, 6, 9, 12 and 18 days post-differentiation (D3-18).

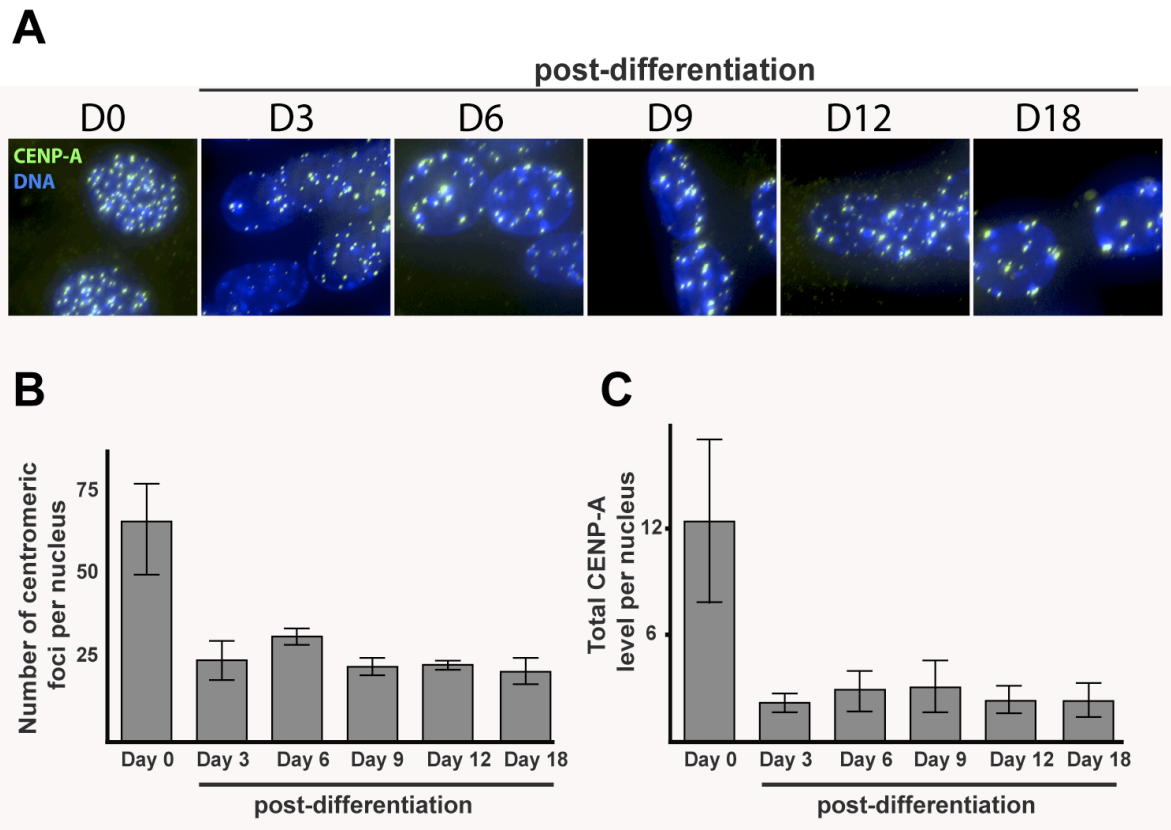

Figure 3 : Centromeres cluster and lose CENP-A upon terminal myogenic differentiation.

A. Nuclei from myoblasts (D0) and terminally differentiated myotubes (D0-D18) stained for CENP-A and DNA (using DAPI) to visualize centromeric foci number, corresponding intensity and subnuclear distribution. B. Average number of centromeric foci per nucleus before (D0) and after (D3-18) differentiation. **C.** Corrected total intensity of CENP-A in nuclei of undifferentiated (D0) and differentiated cells (D3-18).

These observations together suggest that centromeres undergo global changes including overall loss of a significant amount of CENP-A, clustering and spatial rearrangement of the clusters upon exit from the cell cycle (named the differentiated centromere-phenotype, DCP).

Dedifferentiation of myotubes induces spatial redistribution of centromeres:

In order to confirm our findings, we further tested whether the DCP was reversible using the small molecule myosverin, which has been previously demonstrated to induce dedifferentiation of myotubes (Figure 4A-C) (Chang et al., 2001; Ng et al., 2008; Rosania et al., 2000). Accordingly, addition of myoseverin to day 3 post-differentiation myotubes resulted in visible morphological alterations including fragmentation and further individualization of the nuclei (Figure 4A). Upon replacement of the differentiation media with growth media, the individualized cells were observed to restart division as reported previously. Quantification of the number of centromeric signals per nucleus revealed that similar to the above-mentioned observations, the number of centromeric foci dropped dramatically from 51.3 to 15.25 in myoblast-to-myotube transition 3 days post-differentiation (Figure 5A-B). Treatment of day 3 myotubes with myoseverin (followed by incubation in differentiation media), or vehicle alone (followed by incubation in differentiation media containing vehicle) did not result in any significant alteration in the number of centromeric signals with an average number of 13.7 and 19.5 foci per nucleus respectively (Figure 5B). Interestingly, when treated with two consecutive rounds of myoseverin and with myoseverin followed by incubation in growth media, the number of centromeric signals detected underwent an increasing trend from an average number of 27.7 to 53 centromeric foci per

nucleus respectively thus rescuing the DCP in terms of the centromeric signal number (Figure 5B). In addition, cells undergoing various stages of mitosis were observed only in the latter condition.

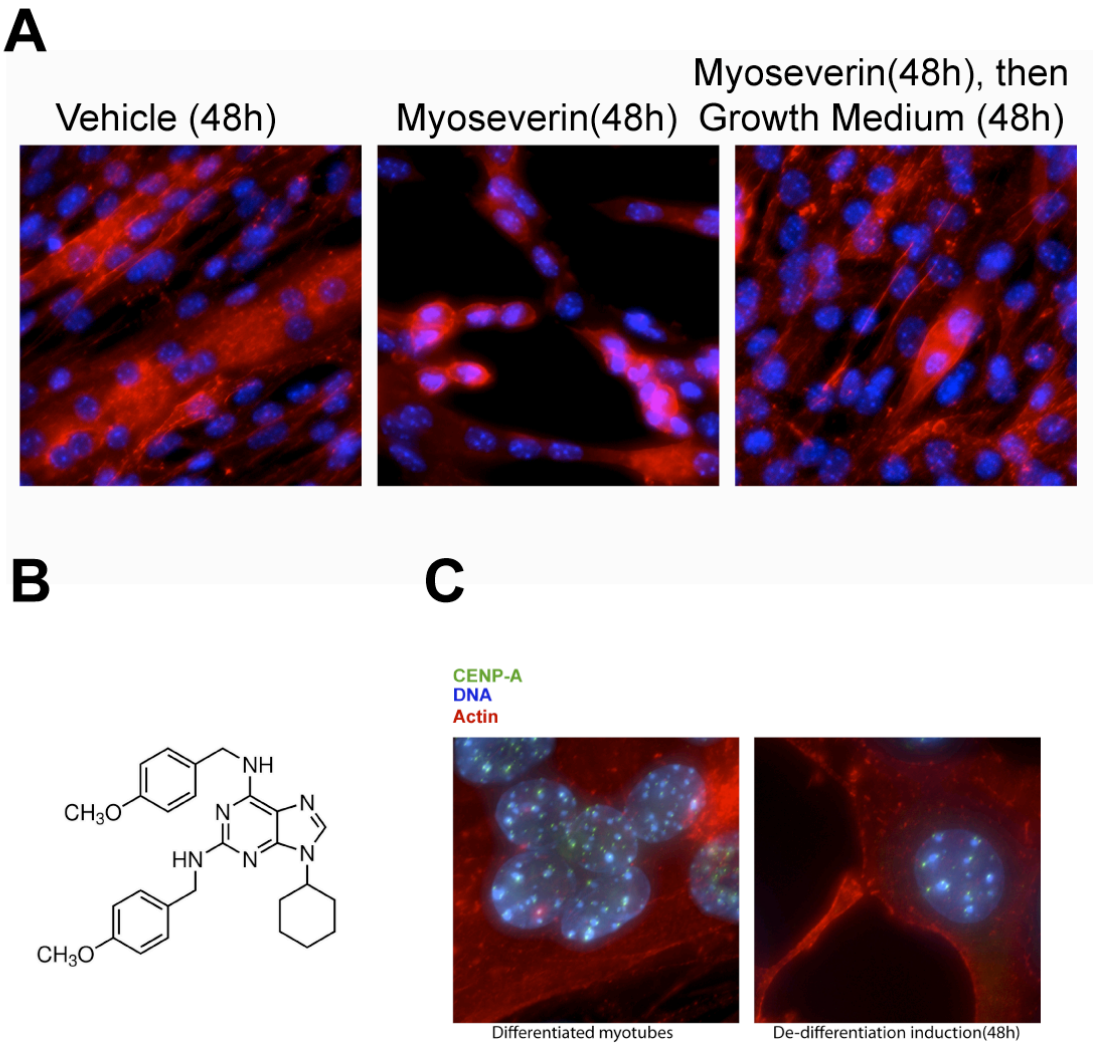

Figure 4 : Myoseverin induces the dedifferentiation of myotubes.

A. Day 3 fully differentiated myotubes (20X) are grown are used for the dedifferentiation experiment. These cells are treated with vehicle for 48 h, fixed and imaged. Myoseverin is added to the differentiation media (DM) and cells are incubated for 48 h. Fragmentation of

myotubes precedes individualization of the nuclei. Upon replacing the media (DM) with proliferation permissive-growth media (GM) cell division is restored in dedifferentiated cells and cell cycle re-entry occurs. **B.** Chemical structure of Myoseverin used in this study. **C.** Zoomed (100X) image of a multi-nucleated differentiated myotube versus a myoseverin-treated cell undergoing dedifferentiation.

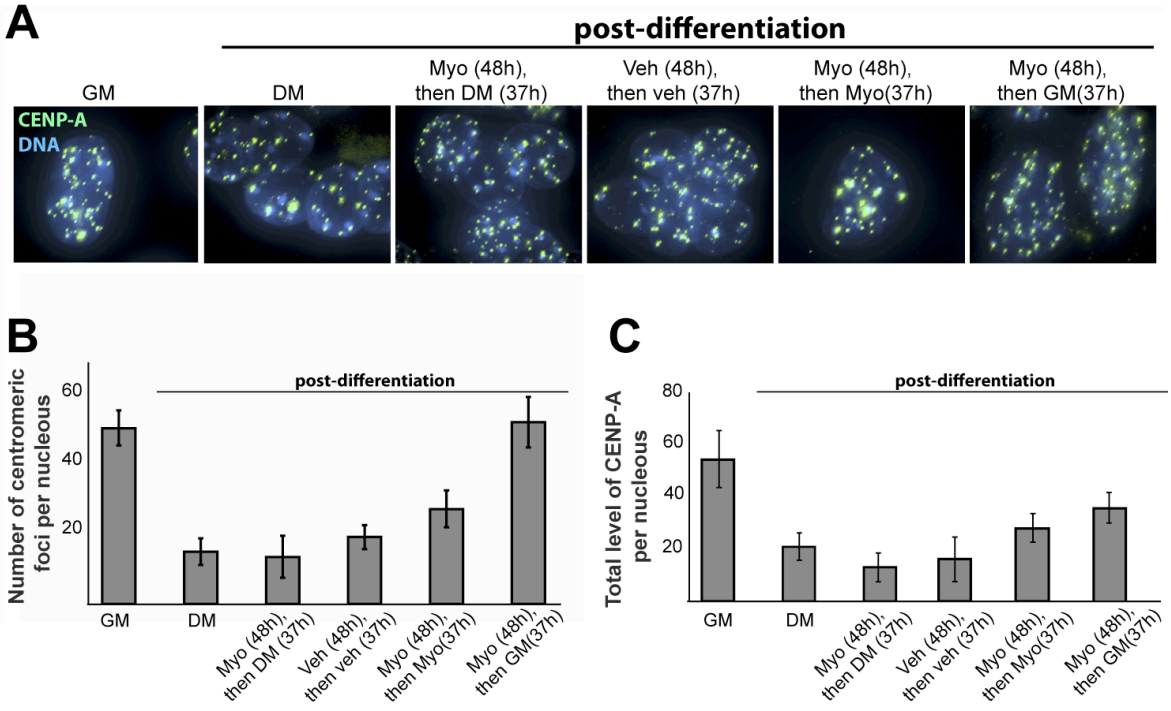

Figure 5 : Centromeres reorganize during dedifferentiation.

A. C2C12 myoblasts were induced to differentiated in DM for 37 h followed by differential treatment with or without myoseverin and incubation in either DM or GM. Cells were fixed and stained for CENP-A and DNA. CENP-A foci number, sub-nuclear distribution and

intensity were measured for quantitative analysis. Note that CENP-A intensity is scaled for visual demonstration. **B.** Average number of CENP-A foci per nucleus were scored in proliferating myoblasts (in GM) and respectively in day 3 differentiated myotubes (in DM), treated with myoseverin (48 h) then incubated in DM (for another 37 h), treated with vehicle (48 h) then incubated in DM containing vehicle (37 h), treated with myoseverin (48 h) then incubated in DM containing myoseverin (37 h) and finally treated with myoseverin (48 h) then incubated in GM (37 h).

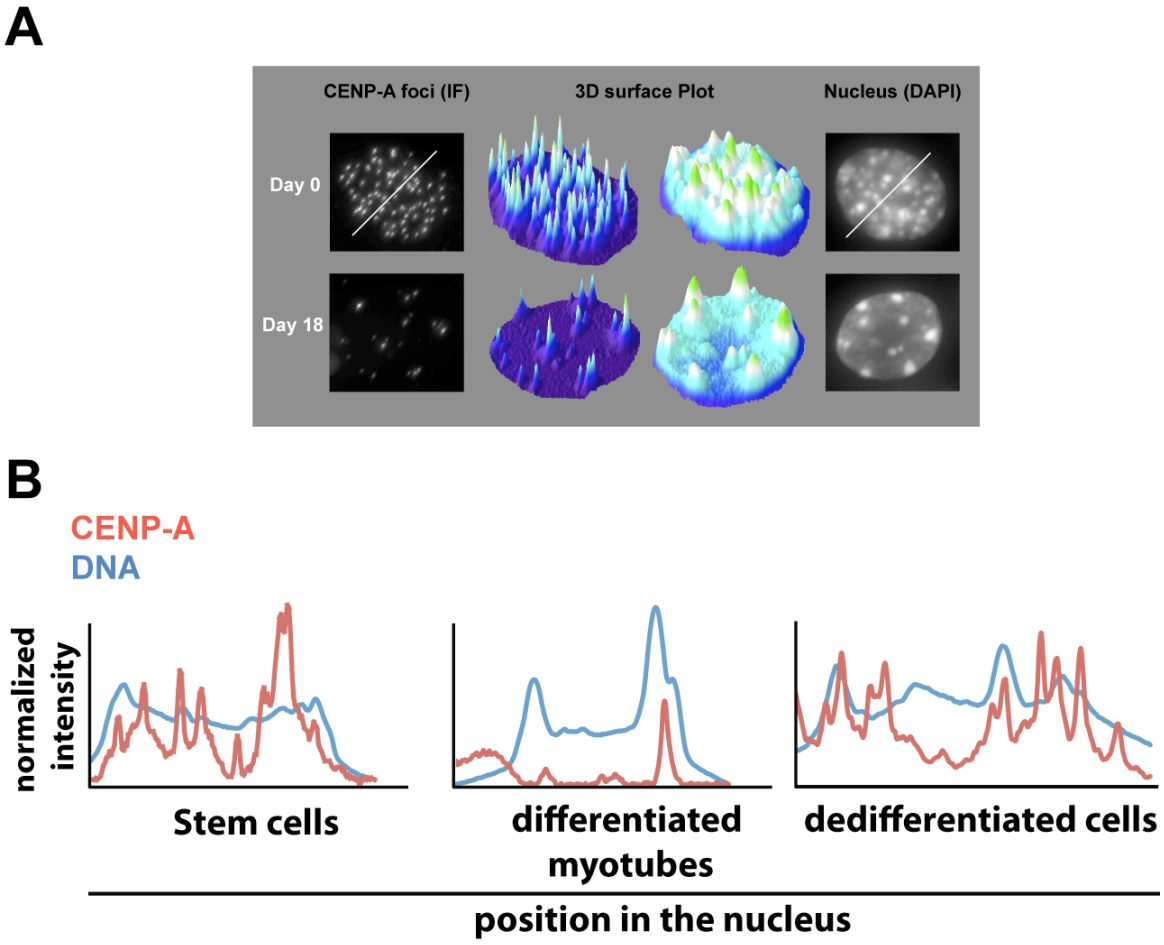

Figure 6 : Centromeres and heterochromatin regions redistribute during differentiation.

A. Line scan analysis of individual nuclei before and after differentiation allows comparison of the spatial distribution of CENP-A foci throughout the nucleus. 3D surface plot of the relative intensities of CENP-A signals confirms our quantification of corrected intensities. **B.** Representative graphs indicate spatial correlation between CENP-A signal sub-nuclear territory and the corresponding DAPI signals. Red and blue peaks represent CENP-A and DAPI-dense heterochromatin foci respectively.

Moreover, line scan analysis of undifferentiated versus differentiated nuclei revealed that in proliferating myoblasts, centromeric signals are detected more frequently as distinct subnuclear entities whereas in day 3 post-differentiation nuclei, centromeric signals almost exclusively colocalized with large heterochromatin foci suggesting the pronounced formation of HECCs (Figure 6A-B). In addition, as exemplified in the 3D surface plot representation (Figure 6A), while in undifferentiated myoblasts, centromeres are evenly distributed throughout the nucleus, in differentiated myotubes higher percentages of centromeres are detected at the nuclear periphery with only 9.8% nuclear periphery-localized centromeres in myoblasts compared to 27.8 %, 23.6 %, 31.4 %, 42.3 % and 36.5 % on days 3, 6, 9, 12, and 18 post-differentiation myotubes respectively (data not shown).

Total CENP-A intensities per nucleus were also quantified for each condition. As anticipated, we observed a marked decrease in total CENP-A intensities upon induction of differentiation (Figure 5C). However, only under conditions when myotubes were induced to dedifferentiate and further incubated with myoseverin for a second pulse, or incubated in growth media following myoseverin treatment, a slight increasing trend was observed reaching a recovery rate of 51.9 % to 65 % respectively (Figure 5C).

Interestingly, upon induction of dedifferentiation and restoration of cell division, centromeric foci appeared to have undergone an overall nuclear repositioning with a nuclear distribution pattern similar to that of undifferentiated nuclei.

Discussion:

In the current work, we have demonstrated that centromeres in wild type mouse cells of various tissue origins may differ significantly in terms of number and total amounts of the

epigenetic mark identifier, CENP-A depending on tissue context (Figure 1A-C). Among tissue sections studied herein, we observed that skin cells contained the most number of centromeres (~25.3 per nucleus) with highest CENP-A levels. Interestingly, images were acquired from the skin epidermis where proliferating stem cells in the basal layer are known to reside (Blanpain and Fuchs, 2006; Fuchs, 2008; Fuchs and Byrne, 1994; Fuchs and Horsley, 2008; Watt, 1998). These stem cells undergo a few rounds of cell division and migrate to the skin surface as they terminally differentiate. The relatively large standard deviation observed for the number of centromeric foci and CENP-A levels in the skin cells could be nicely explained by the fact that our quantifications indeed represent mixed populations of dividing stem cells as well as fully differentiated cells.

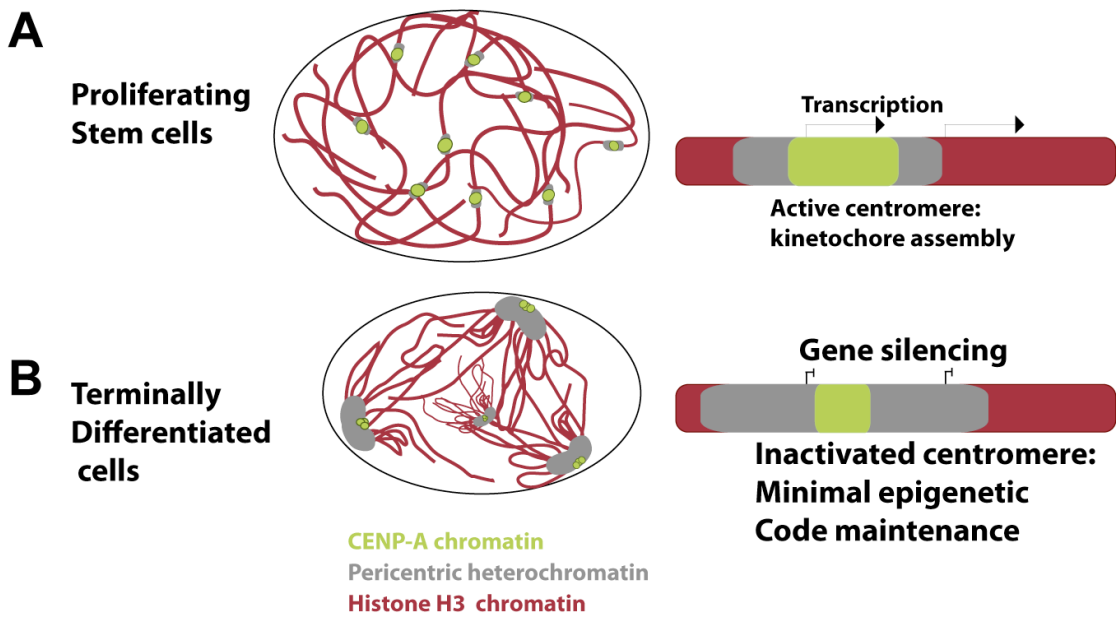

Figure 7 : Centromeric epigenetic mark maintenance and inactivation model.

A. In proliferating cells with an active cell cycle, centromeres are distributed evenly throughout the nucleus. Moreover, CENP-A-containing chromatin occupies a larger portion of

the centric repeats and allows active transcription of centromeric RNA which would in part contribute to mitotic kinetochore assembly. B. In post-mitotic differentiated cells however, centromeres rearrange and coalesce to form centromeric bundles where pericentric heterochromatin expands and eventually invades the formerly CENP-A-occupied region resulting in substantial loss of CENP-A and generation of higher-order chromatin structures. This would ensure that centromeric RNA transcription is repressed, as the cell no longer enters mitosis.

On the other hand, adult heart cells were found to contain the lowest number of centromeric signals with minimal CENP-A intensities. Adult mammalian heart is known to primarily consist of fully differentiated cardiac myocytes, although limited endogenous stem-like cells have also been isolated from this tissue (Leri et al., 2001). The highly differentiated nature of the cardiac myocytes concomitant with the presence of minimal centromeric signals per nucleus could be attributed to the lack of cell division in these cells. Similar observations regarding the centromeric number and CENP-A levels were made for tissues where fully differentiated cells make up bulk of the tissue e.g. brain.

Moreover, we demonstrated that in the C2C12 myogenic model, centromeres undergo dramatic changes upon terminal differentiation culminating in patterns similar to those observed in tissue sections suggesting that centromeres are clustered, repositioned, depleted of an extensive amount of CENP-A and accumulated at heterochromatin sites (Figure 2A-C). As these cells no longer undergo DNA synthesis or mitosis, we hypothesize that these processes occur in order to ensure the maintenance of a minimal centromeric epigenetic code albeit in an inactive state. Our observations of centromeres in dedifferentiated cells, where, under cell

cycle re-entry permissive conditions, centromeric signals increased in number and redistributed throughout the nucleus, provide additional support for our hypothesis.

Intriguingly, the observation of distinctive colocalization of centromeric clusters with large heterochromatin foci (HECCs) may represent the formation of higher order chromatin domains potentially serving versatile purposes in post-mitotic cells (Figure 6A-B). It has been observed that upon differentiation, centromere-derived transcripts are down-regulated (Probst et al., 2010) and the formation of HECCs may in turn contribute to this phenomenon. Centromere RNA has been shown to regulate the activity of aurora B kinase of the chromosomal passenger complex during mitosis in mouse cells (Ferri et al., 2009). Interestingly, aurora B kinase is known to phosphorylate ser7 on CENP-A, an event essential for mitotic kinetochore assembly (Kunitoku et al., 2003; Zeitlin et al., 2001). Accumulation of large centromeric clusters at heterochromatin sites could lead to the spread of pericentric heterochromatin into the CENP-A-containing region, thereby depression of centromeric RNA production, eviction and replacement of CENP-A with other histone H3 variants and thus the overall decreased CENP-A intensity. In addition, the formation of HECCs can contribute to the post-mitotic nuclear architecture by providing higher-order structures functioning as silencer domains for stage-specific regulation of gene expression (Figure 7A-B).

Centromeric loci in higher organisms have been shown to contain repetitive DNA sequences, of which, a fraction is occupied by CENP-A-containing nucleosomes (Vafa and Sullivan, 1997; Warburton et al., 1997). It has also been shown that under normal circumstances, CENP-A exceeds the amount needed to signal the formation of the mitotic kinetochore (Coffman et al., 2011). Most notably, 90% depletion of CENP-A levels does not seem to cause loss of kinetochore activity (and therefore assembly) in HeLa cells (Black et al.,

2007b). Whereas it is unclear why proliferating cells would contain surplus CENP-A, it is plausible that the true epigenetic mark is the minimal inactive form retained in post-mitotic cells. In this scenario, chromosomes would form HECCs in order to inactivate yet conserve the essential elements of the mark throughout the life of the cell.

The mechanisms by which CENP-A is diminished post-differentiation are not known. However, microarray data suggest that the CENP-A transcript level is down-regulated in adult mouse heart compared with embryonic heart (Gupta et al., 2010). Moreover, given that terminal differentiation involves cell cycle exit, and the fact that in the mammalian system, CENP-A is deposited onto the centromeric chromatin in G1(Barnhart et al., 2011; Dunleavy et al., 2009; Foltz et al., 2009; Jansen et al., 2007; Shuaib et al., 2010), the dilution of CENP-A in the last round of S phase could also lead to partial decrease of CENP-A levels in the resultant daughter cells. Action of chromatin remodeling factors by eviction and exchange of CENP-A with other histone H3 variants followed by proteasome-mediated degradation could further contribute to the overall loss of CENP-A levels. Regardless of the underlying molecular players, cells seem to have developed some degree of tolerance to loss of CENP-A. This corroborates very well with our observation where in dedifferentiated cells, while the number of centromeric signals underwent a dramatic increase, total CENP-A intensity showed only a subtle change. Therefore, we anticipate that if these cells are incubated in growth media for longer periods of time, total CENP-A levels would eventually recover to higher degrees.

Altogether, we've shown that centromeres undergo significant remodeling upon terminal differentiation and propose that these changes preserve the minimal epigenetic mark in an inactive state. It would be of great interest to address the composition of CENP-A nucleosomes in terminally differentiated cells using techniques such as total internal reflection

fluorescence microscopy-couple counting of individual molecules previously developed by our group.

Materials and methods:

Cell culture and drug treatment:

Mouse myogenic C2C12 progenitor myoblasts (a kind gift from Craig Mandato, McGill University) were cultured as previously described (Lawson and Purslow, 2000). Briefly, myoblasts were cultured in DMEM supplemented with 10% fetal bovine serum and 1% penicillin/streptomycin (growth medium, GM). In order to induce terminal differentiation and myotube formation, DMEM containing 10% horse serum (differentiation medium, DM) was added to near-confluent myoblasts and cells were incubated for a minimum of 37h or longer periods of time as indicated prior to fixation and staining. Myoseverin was added to day 3 post-differentiation myotubes at 20 μM for 48h. Following myoseverin treatment, cells were further incubated in myosverin-containing or myosverin-free DM or GM as indicated in the text.

Immunostaining and fixed-cell imaging:

C2C12 cells grown on glass coverslips placed at the bottom of petri dishes, were fixed by adding cold (-80 ^{0}C) methanol followed by incubation in -20^{0}C for 20 min. Cells were then washed with PBS 1X three times in intervals of 10 min, and blocked with a PBS1X-TritonX100-BSA 2% solution for 15min at room temperature (RT). Primary antibodies against mouse CENP-A (Cell Signaling Technology) or BrdU (Invitrogen) at 1 mg/ml were diluted 1:1000, added directly to the cell layer followed by incubation for 1 h in the dark at RT. Cells were then washed using PBS1X-TritonX100 solution three times with intervals of 10 min and further incubated with secondary anti-mouse antibodies conjugated to Alexa567 with or without phalloidin 688 diluted 1:250 for 1h in the dark at RT. Excess antibodies were removed

by a final round of wash with PBS 1X. Finally, to visualize the nuclei, mounting medium containing DAPI was added to the interface between the coverslip and a glass slide, followed by sealing of the imaging slide. The slides were then stored in the dark at -20^{0}C until image acquisition.

Imaging was performed as previously described (Lagana et al., 2010) on a DeltaVision widefield microscope equipped with Softworx software (Applied Precision/GE Healthcare) and a CoolSnap2 camera (Photometrics) using 60x or 100x APO plan objectives (Olympus). In order to visualize all centromeric signals, z-stacks were collected for a total thickness of 8 μm from the nuclei. For quantification purposes, Imaging parameters were kept identical in all conditions. Analyses of the resultant images were conducted manually on non-deconvolved images using ImageJ (NIH) software by selecting individual nuclei using the DAPI signal, extracting the fluorescence intensity of centromeric signals in the corresponding channel and finally subtracting the background.

Paraffin-embedded tissue sections were provided from the histology platform tissue library of the Institute for Research in Immunology and Cancer (IRIC) and treated in the same order as described above for immunohistochemistry (IHC) except for primary antibody incubation time (4h).

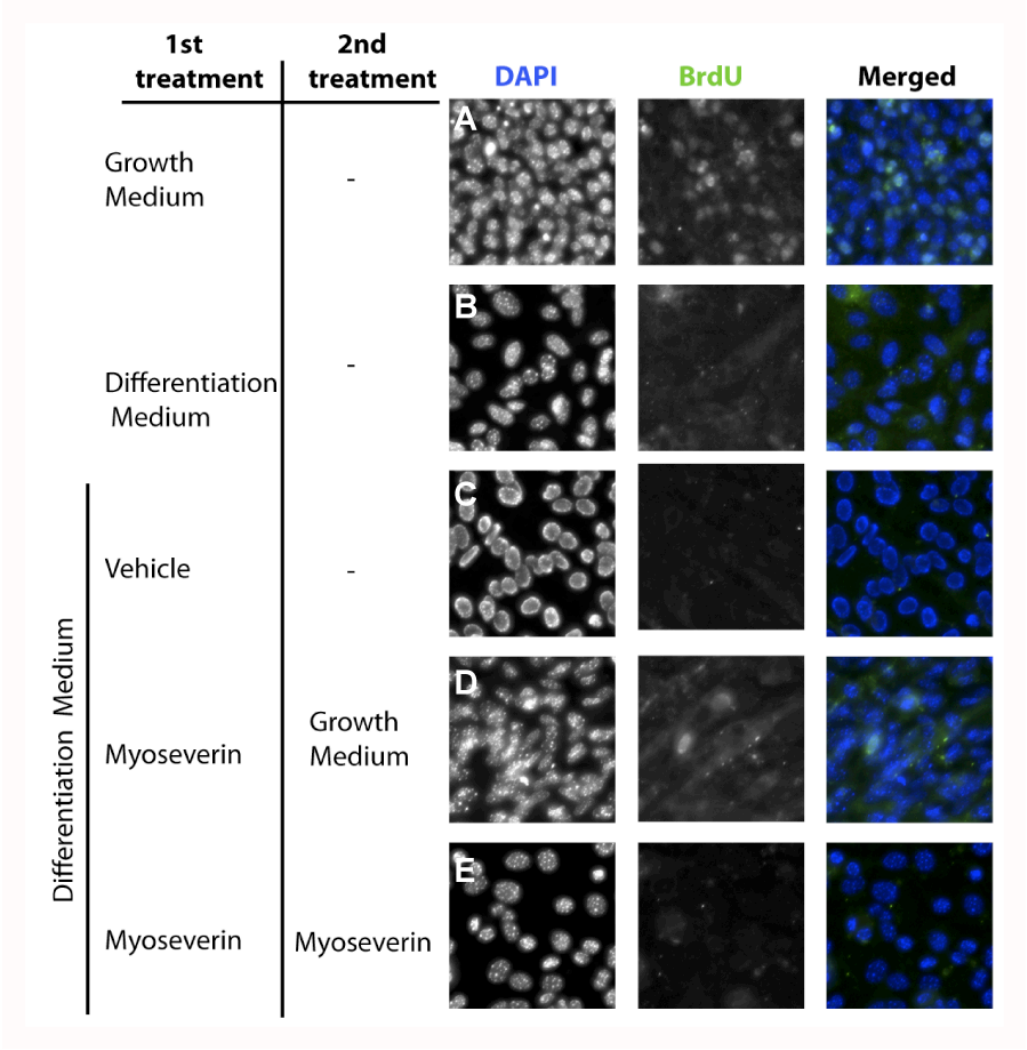

Figure 8 : Supplement I. Dedifferentiated cells can re-enter the cell cycle.

A. BrdU incorporation was used a marker of cycling cells in S phase. As expected, a high number of BrdU positive nuclei was observed in proliferating myoblasts. **B.** In day 3 differentiated myotubes almost no BrdU positive signal was detected indicating that nuclei exit the cell cycle upon terminal differentiation. **C-D.** Day 3 differentiated cells were treated

with either vehicle alone or myoseverin followed by incubation in growth media (GM) or myoseverin-containing differentiation media (DM).

References:

Allshire, R.C., and Karpen, G.H. (2008). Epigenetic regulation of centromeric chromatin: old dogs, new tricks? Nature reviews Genetics *9*, 923-937.

Amato, A., Schillaci, T., Lentini, L., and Di Leonardo, A. (2009). CENPA overexpression promotes genome instability in pRb-depleted human cells. Molecular cancer *8*, 119.

Barnhart, M.C., Kuich, P.H., Stellfox, M.E., Ward, J.A., Bassett, E.A., Black, B.E., and Foltz, D.R. (2011). HJURP is a CENP-A chromatin assembly factor sufficient to form a functional de novo kinetochore. The Journal of cell biology *194*, 229-243.

Black, B.E., Brock, M.A., Bedard, S., Woods, V.L., Jr., and Cleveland, D.W. (2007a). An epigenetic mark generated by the incorporation of CENP-A into centromeric nucleosomes. Proceedings of the National Academy of Sciences of the United States of America *104*, 5008-5013.

Black, B.E., Foltz, D.R., Chakravarthy, S., Luger, K., Woods, V.L., Jr., and Cleveland, D.W. (2004). Structural determinants for generating centromeric chromatin. Nature *430*, 578-582.

Black, B.E., Jansen, L.E., Maddox, P.S., Foltz, D.R., Desai, A.B., Shah, J.V., and Cleveland, D.W. (2007b). Centromere identity maintained by nucleosomes assembled with histone H3 containing the CENP-A targeting domain. Molecular cell *25*, 309-322.

Blanpain, C., and Fuchs, E. (2006). Epidermal stem cells of the skin. Annual review of cell and developmental biology *22*, 339-373.

Bui, M., Dimitriadis, E.K., Hoischen, C., An, E., Quenet, D., Giebe, S., Nita-Lazar, A., Diekmann, S., and Dalal, Y. (2012). Cell-Cycle-Dependent Structural Transitions in the Human CENP-A Nucleosome In Vivo. Cell *150*, 317-326.

Camahort, R., Shivaraju, M., Mattingly, M., Li, B., Nakanishi, S., Zhu, D., Shilatifard, A., Workman, J.L., and Gerton, J.L. (2009). Cse4 is part of an octameric nucleosome in budding yeast. Molecular cell *35*, 794-805.

Chan, F.L., and Wong, L.H. (2012). Transcription in the maintenance of centromere chromatin identity. Nucleic acids research.

Chang, Y.T., Wignall, S.M., Rosania, G.R., Gray, N.S., Hanson, S.R., Su, A.I., Merlie, J., Jr., Moon, H.S., Sangankar, S.B., Perez, O.*, et al.* (2001). Synthesis and biological evaluation of myoseverin derivatives: microtubule assembly inhibitors. Journal of medicinal chemistry *44*, 4497-4500.

Cleveland, D.W., Mao, Y., and Sullivan, K.F. (2003). Centromeres and kinetochores: from epigenetics to mitotic checkpoint signaling. Cell *112*, 407-421.

Coffman, V.C., Wu, P., Parthun, M.R., and Wu, J.Q. (2011). CENP-A exceeds microtubule attachment sites in centromere clusters of both budding and fission yeast. The Journal of cell biology *195*, 563-572.

Dalal, Y., Wang, H., Lindsay, S., and Henikoff, S. (2007). Tetrameric structure of centromeric nucleosomes in interphase Drosophila cells. PLoS Biol *5*, e218.

De Rop, V., Padeganeh, A., and Maddox, P.S. (2012). CENP-A: the key player behind centromere identity, propagation, and kinetochore assembly. Chromosoma.

Dimitriadis, E.K., Weber, C., Gill, R.K., Diekmann, S., and Dalal, Y. (2010). Tetrameric organization of vertebrate centromeric nucleosomes. Proceedings of the National Academy of Sciences of the United States of America *107*, 20317-20322.

Dunleavy, E.M., Almouzni, G., and Karpen, G.H. (2011). H3.3 is deposited at centromeres in S phase as a placeholder for newly assembled CENP-A in G(1) phase. Nucleus *2*, 146-157.

Dunleavy, E.M., Roche, D., Tagami, H., Lacoste, N., Ray-Gallet, D., Nakamura, Y., Daigo, Y., Nakatani, Y., and Almouzni-Pettinotti, G. (2009). HJURP is a cell-cycle-dependent maintenance and deposition factor of CENP-A at centromeres. Cell *137*, 485-497.

Dunleavy, E.M., Zhang, W., and Karpen, G.H. (2013). Solo or doppio: how many CENP-As make a centromeric nucleosome? Nature structural & molecular biology *20*, 648-650.

Ferri, F., Bouzinba-Segard, H., Velasco, G., Hube, F., and Francastel, C. (2009). Non-coding murine centromeric transcripts associate with and potentiate Aurora B kinase. Nucleic acids research *37*, 5071-5080.

Foltz, D.R., Jansen, L.E., Bailey, A.O., Yates, J.R., 3rd, Bassett, E.A., Wood, S., Black, B.E., and Cleveland, D.W. (2009). Centromere-specific assembly of CENP-a nucleosomes is mediated by HJURP. Cell *137*, 472-484.

Fuchs, E. (2008). Skin stem cells: rising to the surface. The Journal of cell biology *180*, 273-284.

Fuchs, E., and Byrne, C. (1994). The epidermis: rising to the surface. Current opinion in genetics & development *4*, 725-736.

Fuchs, E., and Horsley, V. (2008). More than one way to skin. Genes & development *22*, 976-985.

Gupta, M.K., Illich, D.J., Gaarz, A., Matzkies, M., Nguemo, F., Pfannkuche, K., Liang, H., Classen, S., Reppel, M., Schultze, J.L.*, et al.* (2010). Global transcriptional profiles of beating clusters derived from human induced pluripotent stem cells and embryonic stem cells are highly similar. BMC developmental biology *10*, 98.

Heun, P., Erhardt, S., Blower, M.D., Weiss, S., Skora, A.D., and Karpen, G.H. (2006). Mislocalization of the Drosophila centromere-specific histone CID promotes formation of functional ectopic kinetochores. Developmental cell *10*, 303-315.

Jansen, L.E., Black, B.E., Foltz, D.R., and Cleveland, D.W. (2007). Propagation of centromeric chromatin requires exit from mitosis. The Journal of cell biology *176*, 795-805.

Krassovsky, K., Henikoff, J.G., and Henikoff, S. (2012). Tripartite organization of centromeric chromatin in budding yeast. Proceedings of the National Academy of Sciences of the United States of America *109*, 243-248.

Kunitoku, N., Sasayama, T., Marumoto, T., Zhang, D., Honda, S., Kobayashi, O., Hatakeyama, K., Ushio, Y., Saya, H., and Hirota, T. (2003). CENP-A phosphorylation by Aurora-A in prophase is required for enrichment of Aurora-B at inner centromeres and for kinetochore function. Developmental cell *5*, 853-864.

Lagana, A., Dorn, J.F., De Rop, V., Ladouceur, A.M., Maddox, A.S., and Maddox, P.S. (2010). A small GTPase molecular switch regulates epigenetic centromere maintenance by stabilizing newly incorporated CENP-A. Nature cell biology *12*, 1186-1193.

Lange, U.C., Siebert, S., Wossidlo, M., Weiss, T., Ziegler-Birling, C., Walter, J., Torres-Padilla, M.E., Daujat, S., and Schneider, R. (2013). Dissecting the role of H3K64me3 in mouse pericentromeric heterochromatin. Nature communications *4*, 2233.

Lawson, M.A., and Purslow, P.P. (2000). Differentiation of myoblasts in serum-free media: effects of modified media are cell line-specific. Cells, tissues, organs *167*, 130-137.

Leri, A., Quaini, F., Kajstura, J., and Anversa, P. (2001). Myocyte death and myocyte regeneration in the failing human heart. Italian heart journal : official journal of the Italian Federation of Cardiology *2 Suppl 3*, 12S-14S.

Li, Y., Zhu, Z., Zhang, S., Yu, D., Yu, H., Liu, L., Cao, X., Wang, L., Gao, H., and Zhu, M. (2011). ShRNA-targeted centromere protein A inhibits hepatocellular carcinoma growth. PloS one *6*, e17794.

McGovern, S.L., Qi, Y., Pusztai, L., Symmans, W.F., and Buchholz, T.A. (2012). Centromere protein-A, an essential centromere protein, is a prognostic marker for relapse in estrogen receptor-positive breast cancer. Breast cancer research : BCR *14*, R72.

McMahon, D.K., Anderson, P.A., Nassar, R., Bunting, J.B., Saba, Z., Oakeley, A.E., and Malouf, N.N. (1994). C2C12 cells: biophysical, biochemical, and immunocytochemical properties. The American journal of physiology *266*, C1795-1802.

Mizuguchi, G., Xiao, H., Wisniewski, J., Smith, M.M., and Wu, C. (2007). Nonhistone Scm3 and histones CenH3-H4 assemble the core of centromere-specific nucleosomes. Cell *129*, 1153-1164.

Ng, D.C., Gebski, B.L., Grounds, M.D., and Bogoyevitch, M.A. (2008). Myoseverin disrupts sarcomeric organization in myocytes: an effect independent of microtubule assembly inhibition. Cell motility and the cytoskeleton *65*, 40-58.

Padeganeh, A., De Rop, V., and Maddox, P.S. (2013a). Nucleosomal composition at the centromere: a numbers game. Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology *21*, 27-36.

Padeganeh, A., Ryan, J., Boisvert, J., Ladouceur, A.M., Dorn, J.F., and Maddox, P.S. (2013b). Octameric CENP-A nucleosomes are present at human centromeres throughout the cell cycle. Current biology : CB *23*, 764-769.

Probst, A.V., Okamoto, I., Casanova, M., El Marjou, F., Le Baccon, P., and Almouzni, G. (2010). A strand-specific burst in transcription of pericentric satellites is required for chromocenter formation and early mouse development. Developmental cell *19*, 625-638.

Rosania, G.R., Chang, Y.T., Perez, O., Sutherlin, D., Dong, H., Lockhart, D.J., and Schultz, P.G. (2000). Myoseverin, a microtubule-binding molecule with novel cellular effects. Nature biotechnology *18*, 304-308.

Schotta, G., Lachner, M., Sarma, K., Ebert, A., Sengupta, R., Reuter, G., Reinberg, D., and Jenuwein, T. (2004). A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. Genes & development *18*, 1251-1262.

Sekulic, N., Bassett, E.A., Rogers, D.J., and Black, B.E. (2010). The structure of (CENP-A-H4)(2) reveals physical features that mark centromeres. Nature *467*, 347-351.

Shuaib, M., Ouararhni, K., Dimitrov, S., and Hamiche, A. (2010). HJURP binds CENP-A via a highly conserved N-terminal domain and mediates its deposition at

centromeres. Proceedings of the National Academy of Sciences of the United States of America *107*, 1349-1354.

Sullivan, B.A., and Karpen, G.H. (2004). Centromeric chromatin exhibits a histone modification pattern that is distinct from both euchromatin and heterochromatin. Nature structural & molecular biology *11*, 1076-1083.

Tachiwana, H., Kagawa, W., Shiga, T., Osakabe, A., Miya, Y., Saito, K., Hayashi-Takanaka, Y., Oda, T., Sato, M., Park, S.Y.*, et al.* (2011). Crystal structure of the human centromeric nucleosome containing CENP-A. Nature *476*, 232-235.

Tomonaga, T., Matsushita, K., Yamaguchi, S., Oohashi, T., Shimada, H., Ochiai, T., Yoda, K., and Nomura, F. (2003). Overexpression and mistargeting of centromere protein-A in human primary colorectal cancer. Cancer research *63*, 3511-3516.

Vafa, O., and Sullivan, K.F. (1997). Chromatin containing CENP-A and alpha-satellite DNA is a major component of the inner kinetochore plate. Current biology : CB *7*, 897-900.

Valdivia, M.M., Hamdouch, K., Ortiz, M., and Astola, A. (2009). CENPA a genomic marker for centromere activity and human diseases. Current genomics *10*, 326-335.

Van Hooser, A.A., Ouspenski, II, Gregson, H.C., Starr, D.A., Yen, T.J., Goldberg, M.L., Yokomori, K., Earnshaw, W.C., Sullivan, K.F., and Brinkley, B.R. (2001). Specification of kinetochore-forming chromatin by the histone H3 variant CENP-A. Journal of cell science *114*, 3529-3542.

Warburton, P.E., Cooke, C.A., Bourassa, S., Vafa, O., Sullivan, B.A., Stetten, G., Gimelli, G., Warburton, D., Tyler-Smith, C., Sullivan, K.F.*, et al.* (1997). Immunolocalization

of CENP-A suggests a distinct nucleosome structure at the inner kinetochore plate of active centromeres. Current biology : CB *7*, 901-904.

Watt, F.M. (1998). Epidermal stem cells: markers, patterning and the control of stem cell fate. Philosophical transactions of the Royal Society of London Series B, Biological sciences *353*, 831-837.

Yoda, K., Ando, S., Morishita, S., Houmura, K., Hashimoto, K., Takeyasu, K., and Okazaki, T. (2000). Human centromere protein A (CENP-A) can replace histone H3 in nucleosome reconstitution in vitro. Proceedings of the National Academy of Sciences of the United States of America *97*, 7266-7271.

Zeitlin, S.G., Shelby, R.D., and Sullivan, K.F. (2001). CENP-A is phosphorylated by Aurora B kinase and plays an unexpected role in completion of cytokinesis. The Journal of cell biology *155*, 1147-1157.

Zhou, Z., Feng, H., Zhou, B.R., Ghirlando, R., Hu, K., Zwolak, A., Miller Jenkins, L.M., Xiao, H., Tjandra, N., Wu, C.*, et al.* (2011). Structural basis for recognition of centromere histone variant CenH3 by the chaperone Scm3. Nature *472*, 234-237.

Chapter 4

General discussion, conclusion and perspectives:

CENP-A, a common feature of centromeres throughout evolution

Centromeric loci, in most species are composed of repetitive DNA and vary in size, sequence and physical locations even from chromosome to chromosome in the same organism (Sullivan et al., 2011; Zhang and Dawe, 2012). Regional centromeres in humans range in size from 1-15 Mbp, while point centromeres in *Saccharomyces cerevisiae* occupy only 125 bp of DNA (Smith, 2002; Torras-Llort et al., 2009; Verdaasdonk and Bloom, 2011) and diffuse/holocentric centromeres in *Caenorhabditis elegans* form axes all along the chromosome arms (Dernburg, 2001; Maddox et al., 2004; Monen et al., 2005). Figure 1 summarizes some of the key differences in centromeres of various organisms. In most cases (except for budding yeast), however, the underlying DNA sequence has been shown neither to be necessary nor essential for the specification of centromeres (Craig et al., 1999; Henikoff et al., 2001; Sullivan et al., 2001). Therefore, based on several lines of evidence, an epigenetic mechanism has been proposed to play a key role in determining centromeric identity (Cleveland et al., 2003; Karpen and Allshire, 1997). One of the key common features of all centromeres throughout evolution regardless of mechanism(s) of specification is the presence of the histone H3 variant, centromere protein-A, CENP-A, at the centromeric chromatin (Cleveland et al., 2003). Homologues of CENP-A (originally given this name in the mammalian system) have been identified and extensively studied in other organisms. CENP-A, in *Caenorhabditis elegans* called HCP3 (HoloCentric chromosome binding Protein), in Saccharomyces *cerevisiae* referred to as CSE4 (Chromosome SEgregation-4), in Arabidopsis thaliana called CENH3 (CENtromere-specific histone H3),

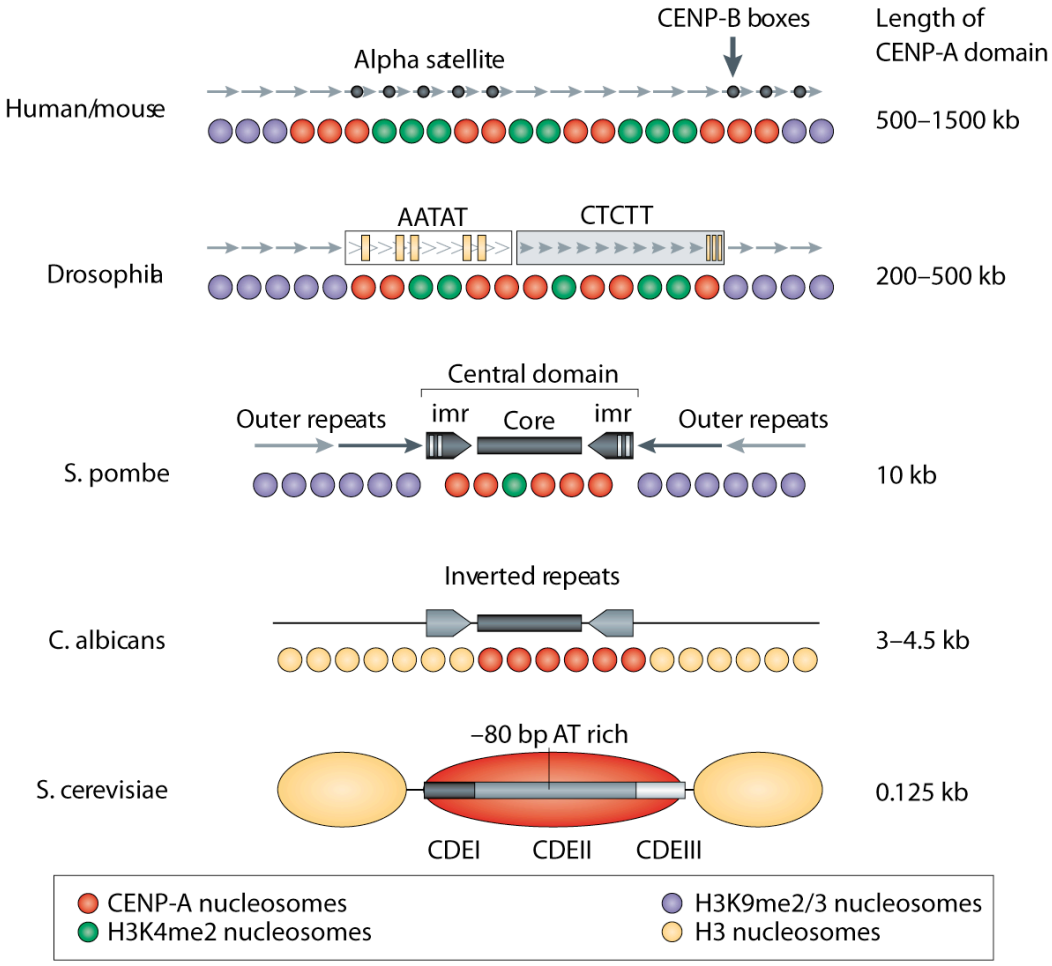

Figure 1 : Centromeres vary in size and DNA sequence throughout evolution.

Presence of CENP-A-containing nucleosomes is a common feature among all centromeres in different organisms whereas their sizes range from ~125 bp in budding yeast to kilo bases and Mega bases of repetitive sequences in higher species e.g. drosophila and humans.

in Drosophila melanogaster known as CID (Centromere IDentifier) and in Schizosaccharomyces pombe called Cnp1, is incorporated at the centromeric chromatin in the form of nucleosomes providing a structural foundation for the assembly of the mitotic kinetochore (Cleveland et al., 2003). Therefore, the composition of CENP-A-containing nucleosomes determines the structural nature and potentially mechanism of action of the CENP-A epigenetic code. Given that the same code is read and translated differently at critical time points of the cell cycle, we reasoned that knowing the molecular make-up of the CENP-A-containing nucleosomes would hold the key to understanding the differential CENP-A-mediated roles in the cell.

Development of the single-molecule imaging assay:

In order to reduce and possibly eliminate some of the caveats about other approaches e.g. *in vitro* nucleosome reconstitution (Tachiwana et al., 2011a; Yoda et al., 2000), atomic force microscopy (Bui et al., 2012; Dalal et al., 2007; Dimitriadis et al., 2010) and fluorescence correlation spectroscopy (Shivaraju et al., 2012), there needed to exist a technique that would allow us to directly count the copy number of CENP-A molecules in "in-cell-assembled " nucleosomes.

Therefore, we set to develop a novel molecular counting assay based on total internal reflection fluorescence microscopy (TIRFm). This assay, not only allows us to directly visualize (with minimal noise) single-molecules for qualitative analyses e.g. protein-protein interactions and the spatial localization of two or more proteins, it also enables us to gain quantitative information about the copy number of a given protein in a given protein complex. Interestingly, this, and other versions of this approach (referred to single-molecule pull down

by others) (Jain et al., 2011; Jain et al., 2012; Padeganeh et al., 2013) are amenable to a wide range of molecular studies far beyond fundamental biology and easily expandable to other areas of research such as medicine. In this regard, our TIRF-based assay could be used to generate molecular biosensors to screen for novel small molecule inhibitors targeting the dimerization, and thereby activation of frequently mutated or over-expressed oncogenes, e.g. human epidermal growth factor receptor 2 (HER2) in cancer (Gordon et al., 2006). We envision that this assay's versatility and precision will be a very suitable complementary approach for emerging drug-screening techniques such as bioluminescence resonance energy transfer (*BRET*) (Lavoie et al., 2013).

Octameric CENP-A nucleosomes reside at the centromeric chromatin:

Previous work had resulted in controversial observations regarding the stoichiometry of histones in the centromeric nucleosome. Perhaps the most striking differences are between the octameric nucleosome model and the hemisome/half-nucleosome model. Atomic force microscopy-based analyses indicated that cross-linked CENP-A/CID nucleosomes obtained from cultured human or drosophila cells had half the height of canonical H3 nucleosomes. This led to the proposal of the hemisome/half-nucleosome model, where heterotypic (CENP-A:H4:H2B:H2A) tetramers are suggested to be found at the centromeric chromatin (Bui et al., 2012; Dalal et al., 2007; Dimitriadis et al., 2010). On the contrary, *in-vitro* reconstituted CENP-A nucleosomes using purified histones have been reported to exist as octamers (Tachiwana et al., 2011a; Yoda et al., 2000) and until recently, CENP-A hemisomes had not been reproduced *in-vitro*. However, lately in an attempt to reconcile the lack of *in-vitro* support, Cse4 hemisomes were reported to have been generated using budding yeast proteins

and an alternative reconstitution protocol (Furuyama et al., 2013). The authors argue that the standard protocols commonly used for nucleosome reconstituting use 2 M NaCl, 147 bp DNA fragments of highly acidic nature and basic histones, which favor the formation of octamers. Their data suggest that that Cse4 octamers spontaneously assembled upon co-expression of all core histones in bacteria, were much more resistant to low salt denaturation (0.5 M NaCl) and release of H2B:H2A dimers compared to canonical H3 octamers thereby introducing more bias into conclusions. Therefore, the authors, instead, employed two 62 bp DNA duplexes each wrapping around a hemisome to produce "pseudo-octamers" and exposed them to low salt dialysis. Using native gel-shift assays they show that Cse4 forms a single band whereas canonical H3 produces more than one band on the gel. AFM analysis of the resultant complexes suggests that Cse4 particles are ~2 nm high (proposed by the authors to be Cse4 hemisomes) while H3 particles have two distinct populations with ~2.2 nm and ~3-5 nm heights (proposed to be H3 hemisomes and octameric aggregates respectively). Moreover, they show that when the DNA duplex ends are fluorescently labeled with FRET donors and acceptors, Cse4 hemisomes not only show 10-15% FRET but also remain stable at high 4-8 M urea suggesting tight wrapping and high stability of such complexes (Furuyama et al., 2013). However, in this approach, since duplexes wrap around individual hemisomes, they generate a gap between the dyad axes in the resultant "pseudo-octamer" which may significantly alter the stability of the octamer and biophysically differ from true octameric nucleosome assembled in cells. This is best exemplified in their observation of H3 hemisomes, not reported to be have been produced using other techniques, compromising the *in-vivo* relevance of these findings.

On the other hand, another group has recently reevaluated these observations by performing AFM measurements on *in-vitro* assembled octameric CENP-A nucleosomes using

human or *Schizosaccharomyces pombe* CENP-A. Surprisingly, their data indicate that octameric CENP-A nucleosomes assembled from human and *S. pombe* proteins have median heights of ~1.64 nm and ~ 0.96 nm, which are consistently lower than that of canonical H3 octameric nucleosomes i.e. 2.09 nm and 1.43 nm respectively (Miell et al., 2013; Miell et al., 2014). After ruling out that the reported height differences are not an artifact of sample preparation, variable imaging conditions or nucleosome deformation, the authors generate human $H3^{CATD}$ hybrid molecules (where H3 histone fold domain loop 1 and alpha helix 2 are substituted with that of CENP-A) and assemble octameric nucleosomes. As mentioned in chapter 1, loop 1 and alpha helix 2 in CENP-A (collectively referred to as the CENP-A targeting domain, CATD) are sufficient to specifically target this protein to the centromere. Interestingly, AFM measurements reveal that octameric $H3^{CATD}$ nucleosomes have a median height of 1.43 nm, which is lower than that of 2.09 nm in H3 octamers. In sum, the authors conclude that CENP-A octamers are 21-33% shorter than H3 octamers. In addition, the authors observe that when subjected to micrococcal nuclease digestion (MNase), *in-vitro* assembled octameric CENP-A nucleosomes used in AFM measurments, protect 20 bp less DNA than H3 nucleosomes suggesting partial unwrapping of DNA at nucleosome entry/exit sites consistent with previous data (Miell et al., 2013).

However, shortly after the publication of these observations, conflicting data from two other reports were presented (Codomo et al., 2014; Walkiewicz et al., 2014). In the first paper (Codomo et al., 2014), the authors prepared Cse4 hemisomes and octasomes *in-vitro* using short and long fragments of α-satellite DNA dialyzing them through 4 M urea. Using AFM measurements they observed that while Cse4 and H3 octasomes displayed similar heights, their corresponding hemisomes were 35-44% shorter. In the second paper (Walkiewicz et al.,

2014), human CENP-A or Cse4 were used to reconstitute octameric nucleosomes and their heights were compared with reconstituted H3 nucleosomes by AFM. The authors however did not detect any significant difference between the heights of CENP-A and H3 octamers. It is therefore argued that the previous observation that CENP-A octamers confer a reduction in height might be the result of artifacts such as the release of H2A:H2B dimers or hydration level. Additionally, AFM measurements are subject to high variability from group to group depending on experimental conditions (Miell et al., 2014) and therefore should be handled and interpreted with much care. As a result, in order to draw any conclusion about stoichiometry of protein complexes, in this case CENP-A nucleosomes, AFM-derived data must be used in combination with other complementary approaches. Interestingly, in agreement with the octamer model, another paper recently described important properties of human CENP-A nucleosomes isolated from cancer cells or cell lines harboring a functional neocentromere (Hasson et al., 2013). MNase analysis of CENP-A-ChIP products revealed that CENP-A particles generate three discreet nuclease-protected DNA fragments with lengths of ~ 110 bp, ~ 130 and ~ 150 bp, different from the single ~165 bp-protected fragment obtained from bulk chromatin. Interestingly, heavy digestion of CENP-A particles originating from centromeres and neocentromeres, eliminated the larger fragments and persistently resulted in a single ~110 bp fragment suggesting that CENP-A nucleosomes have partially unwrapped termini further digested by extensive MNase treatment. Moreover, the authors show that CENP-A nucleosomes of the three sizes reported belong to the same α-satellite or complex DNA positions in normal or neocentromere-containing cells and importantly that the phasing of CENP-A and bulk nucleosomes are very similar in normal centromeres at α-sattelite sequences. Finally, the authors show that regardless of the origin of the centromere, DNA

termini unwrapping seems to be a common feature for CENP-A nucleosomes. Altogether, these observations provided support for the octameric form of CENP-A nucleosomes and suggested that partial unwrapping of the DNA ends are conserved physical characteristics of such nucleosomes.

In this thesis project, considering these extensive efforts, tackling the controversy was inevitable. However, our single-molecule approach uses minimally manipulated naively assembled CENP-A nucleosomes and allows direct visualization and stoichiometric analyses of the copy number of CENP-A per particle. Accordingly, the detection of octameric CENP-A nucleosomes at human centromeres regardless of the cell cycle (Dunleavy et al., 2013; Padeganeh et al., 2013) indicates that cellular components reading and interacting directly with the CENP-A epigenetic code at different time points of the cell cycle recognize the same basic structure. Therefore, additional mechanisms should exist to ensure that the code is not misinterpreted at a given time.

A lot of key questions remain to be answered; it would be of great interest to determine how and at which stoichiometry centromere-associated proteins such as CENP-B, CENP-C or CENP-T bind to the CENP-A nucleosome and how they lead to the formation of the higher order kinetochore in a mitosis-restricted manner. In addition, it is currently unclear how the gaps generated as the result of the dilution of the preexisting CENP-A nucleosomes in S phase are temporarily filled. To answer this question, integration of real-time chromatin replication to our TIRFm assay would be an ideal approach. Identification of other factors or chaperone proteins mediating centromeric chromatin specification, assembly and inheritance from one division to the next is a critical step. One important improvement of our assay to answer some

of these questions is to use simpler model organisms e.g. budding yeast or drosophila where addition or deletion of a given protein is relatively easily achievable. For example, replacement of the endogenous copies of multiple proteins and generation of viable cells co-expressing multiple fluorescently labeled centromere-associated proteins would allow us to simultaneously assess the copy number of our proteins of interest associated to individual CENP-A nucleosomes.

In spite of the above-mentioned lines of evidence in support of the presence of octameric CENP-A nucleosomes throughout the cell cycle, there remains a possibility that a subtle population of CENP-A would have the monomeric form at the centromeric chromatin. , This possibility could not be fully excluded at the present particularly given our observation of a slight decrease in the dimeric form of CENP-A in cell cycle-staged chromatin. Several scenarios could be used to explain the presence of cemtromere-derived monomeric CENP-A: it is presumable that a proportion of dimeric CENP-A is split into monomeric CENP-A during S-phase and remains as such throughout the rest of the cell cycle until next G1 where newly synthesized CENP-A. This hypothesis is consistent with the fact that the decrease in the dimer count was observed to be more pronounced during early S, late S, G2 and M compared to G1. In this regard, it is also possible that the monomeric CENP-A actually dimerizes with histone H3 variant H3.3 to form a chimeric octamer, $(CENP\text{-}A{:}H3.3)_2(H2A{:}H2B)_2$ acting as a placeholder within the gaps generated during S phase for CENP-A. Another possibility is that the hemisome model would indeed exist, albeit as a minority population at the centromeres regardless of the cell cycle status. In our hands however, this population appears to be under-represented hence termed the minority population whereas the major form detected was dimeric CENP-A.

On the other hand, it is also very likely that the monomeric CENP-A detected represents assembly intermediates originating from the non-chromatin CENP-A pool or cytosolic CENP-A. As indicated, around 50% of CENP-A associated to HJURP is the monomeric form. Since we performed the molecular counting assay on asynchronized chromatin and CENP-A expression is known to begin from S to G2 phase, it is very likely that the monomeric form is in fact the HJURP bound form of CENP-A. However, this composition is not prone to deposition until G1 following the action of the licensing complex, which renders the centromeric chromatin receptive to CENP-A assembly. In order to test these hypothetical models, several approaches could be employed. For example, using cells expressing CENP-A-YFP, chromatin stretching followed by immunostaining against H3.3 could be used. It is anticipated that if CENP-A forms a chimeric nucleosomal composition, we would detect the co-localization of CENP-A with H3.3 within interspersed regions of the CENP-A chromatin. In addition, to determin whether the monomeric form detected is nucleosomal or non-nucleosomal form; immuno-staining against HJURP could be used in stretched chromatin preparation. We anticipate detecting no co-localization between HJURP and CENP-A at the stretched centromeric chromatin. However, if we detect co-localization, this could indicate that monomeric CENP-A bound to HJURP could be the source of our monomer count. Another approach to address this issue is to use an *in-vitro* chromatin replication assay where fluorescently labeled histones e.g. CENP-A-YFP and histone H3.3-RFP are visualized by TIRF imaging while undergoing replication. The dynamics of CENP-A dilution to daughter strands as well as stochiometric properties could be directly assessed. By increasing the laser power post-replication the counting the photo-bleaching steps of resultant

particles, it would be possible to determine whether monomeric CENP-A is generated as half nucleosomes (not colocalizing with H3.3) or forms a chimeric nucleosome with H3.3.

Proposed evolutionary pathway for Octameric CENP-A nucleosomes:

Our data indicate that CENP-A is incorporated at the centric chromatin in the form of homotypic octameric nucleosomes (CENP-A:H4:H2B:H2A)$_2$. On the other hand, other core histone H3 variants e.g canonical H3.1 or H3.3 have also been shown to exist in octameric nucleosomes throughout the genome (Jin and Felsenfeld, 2007; Tachiwana et al., 2011b). However, as mentioned before, archeal nucleosomes are known to have a homo-tetrameric nature (H3:H4)$_2$(Pereira et al., 1997; Pereira and Reeve, 1998). This immediately raises the question of how octameric CENP-A nucleosomes containing H2A:H2B dimers could have evolved from ancestral homotetramers. One possible evolutionary scenario is that ancestral (H3:H4)$_2$ tetramers could have gone through extensive divergence to give rise to (H2B:H2A) dimers temporarily coexisting with (H3:H4) dimers to form intermediate (H3:H4:H2B:H2A) heterotetramers. This, presumably unstable complex, further rapidly diverged to produce an independent entity for (H2B:H2A) dimers. According to this model, at this point, these core histones had to undergo massive point mutations e.g. in H3 histone fold domain (HFD) particularly the L1 loop and alpha 2 helix in order to generate a specialized histone H3-variant with a CENP-A targeting domain (CATD), and change histone H4 N- and C-terminal Helix-Sheet-Helix (HSH) domains to accommodate (H2B:H2A) binding while allowing the dimerization of two (CENP-A:H4) complexes and eventual octameric nucleosome formation.

CENP-A nucleosomes and chromatin in post-mitotic cells:

In spite of extensive work on CENP-A protein, CENP-A-containing complexes, CENP-A role in mitotic kinetochore assembly and the epigenetic mark conferred by CENP-A, the dynamics of CENP-A outside the cell cycle e.g. in different developmental stages are poorly known. In a pioneering study, it was demonstrated that while heterozygous mice for CENP-A gene disruption appear healthy, homozygous null mutant mice for cenp-a are embryonic lethal at 6.5 days post-conception (Howman et al., 2000). This lethality was accompanied by a whole range of aberrant phenotypes including the lack of a defined inner cell mass, and cellular incoherency as well as sever mitotic defects e.g. chromosome missegregation, micronuclei formation, chromosome lagging, failure of proper cytokinesis and nuclear membrane blebbing (Figure 2) (Howman et al., 2000). Therefore, CENP-A is an essential gene for embryonic development and viability. One of the key features of development is cellular differentiation. CENP-A dynamics through stem cell differentiation remains elusive. In an attempt to address this question, the Clark lab used human Embryonic Stem Cells (hESCs) and human Induced Pluripotent Stem cells (hIPS) and compared CENP-A levels with a primary skin fibroblast cell line (Ambartsumyan et al., 2010). They observed that CENP-A mRNA levels were 10-fold higher in hESCs and hIPS cells relative to that of fibroblasts. However, surprisingly, they report that CENP-A protein levels measured by western blotting were equal between undifferentiated pluripotent stem cells (hPSCs) and fibroblasts. Moreover, they observed that upon differentiation induction of hPSCs in the

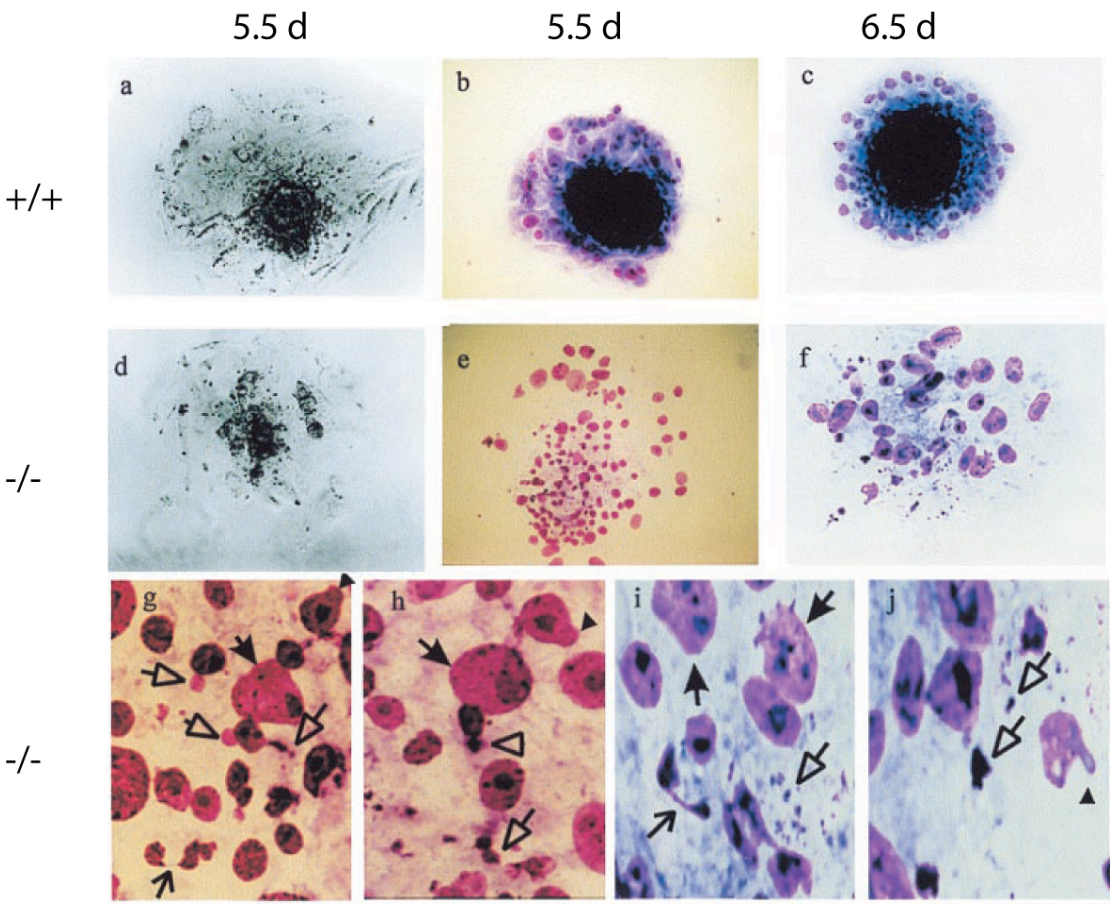

Figure 2 : Cenp-a gene null mutant mice are embryonic lethal and display various types of mitotic errors.

While type embryos imaged using: **A.** phase contrast, **B.** Giemsa staining on day 5.5 and **C.** Giemsa staining of day 6.5 embryo. **D-F.** Counterpart images of **A-C** from cenp-a null mutant embryos. **G** and **H** are magnifications of **E. I** and **J** are magnifications of **F.** Filled arrows indicate macro-nuclei, empty arrows indicate micro-nuclei, open arrows indicate nuclear bridging and finally filled arrowheads represent nuclear blebbing (Howman et al., 2000).

presence of retinoic acid for 7 days, mRNA levels of CENP-A do not undergo any change while in differentiated cells CENP-A levels are increased (Ambartsumyan et al., 2010). This is contrary to our observations in C2C12 myoblasts differentiation where CENP-A protein levels are observed to significantly decrease upon terminal differentiation. This discrepancy could be due to the biological nature e.g pluripotency degree or technical issues such as inefficient protein extraction or unequal loading on the gel. Moreover, their data indicated that shRNA depletion of CENP-A in hPSCs did not affect any of pluripotency-related properties such as cellular morphology, apoptosis and proliferation rates or self-renewal markers expression.

Our data indicate that in terminally differentiated cells, centromere nuclear distribution and CENP-A protein levels differ from that of proliferating stem cells with an active cell cycle. Here we reason that cellular processes such as loss of extensive amounts of CENP-A protein along with bundling of centromeres would trap the minimal epigenetic mark in a dormant state. In addition, the association of centromeres with large heterochromatin structures would further occlude binding of the molecular machinery responsible for the reading and translation of the epigenetic code.

Extensive loss of CENP-A in post mitotic cells as shown by weaker CENP-A immunostaining in fully differentiated myotubes and confirmed by the observation of minimal amounts of CENP-A in adult mouse tissue sections could be explained by a number of potential reasons. As mentioned previously, CENP-A transcript levels have been shown to be down regulated in iPS-derived beating clusters. This decrease has been reported to be more pronounced in the adult heart indicating that CENP-A expression is down regulated globally at the mRNA level to reduce the production of nascent CENP-A. Moreover, ubiquitin-dependent

proteolysis could further contribute to the removal of the already assembled CENP-A at the centromeric chromatin. It is possible that the proportion of CENP-A removed is no longer needed by the cell to maintain the inactive minimal epigenetic mark and therefore cells have evolutionarily developed mechanisms to avoid consuming extra energy to produce a cellular product that no longer serves a purpose.

Here we propose that the expansion of heterochromatin into the CENP-A containing region not only contributes to the inactivation of the CENP-A epigenetic mark, but also it could potentially affect the nucleosome composition in this area. In this model, other DNA synthesis-independent histone H3 variants e.g. H3.3 could, via a step-wise substitutive process, replace CENP-A and convert the active centromere identity to a silent heterochromatin nature. This would involve the action of unknown chromatin remodeling factors mediating the release of H2A:H2B dimers from the CENP-A octamer initially followed by the substitution of CENP-A or (CENP-A:H4) with H3.3-containing counterparts and HP1 binding to render the heterochromatin properties to the resultant chromatin. This model is summarized in Figure 3.

The exact role of CENP-A in post-mitotic cells is not known to date. However, we propose that this proportion while maintaining the minimal epigenetic mark, plays a structural in non-proliferating cells by signaling the generation of sub-nuclear sites where pericentric heterochromatin can accumulate and bind to the nuclear periphery. These regions, including clustered centromeres within expanded heterochromatin, would then act as “hot spots” for the reorganization of the bulk chromatin to silence the expression of certain genes e.g. self-renewal or pluripotency genes. Therefore, we hypothesize that the minimal CENP-A

transcription and protein presence at the basal level is essential for this structural role. To test this model, a variety of systems could be used. These include transgenic stem cells harboring a conditional LoxP-induced knockout system for CENP-A gene targeting or C2C12 myoblasts transduced with a lentiviral plasmid encoding a shRNA against CENP-A under a drug-induced promoter. These stem cells could then be incubated in the differentiation medium to allow cell cycle exit and the formation different cell types. After confirming the clustering of the centromeres at heterochromatin regions, CENP-A gene deletion or shRNA production would be induced to fully deplete the residual transcription of the CENP-A gene. If our hypothesis is correct, we anticipate observing the deformation of the post-mitotic subnuclear structure evident by an increase in the number of remaining centromere foci and the declustering of the formerly detected large heterochromatin foci. In order to test if post-mitotic CENP-A along with its associated heterochromatin influence gene expression, transcriptome-wide microarray could be conducted following complete depletion of residual CENP-A transcripts. We expect to observe the upregulation of cell lineage non-specific transcripts in these cells.

Moreover, it is intuitive to assume that the minimally retained CENP-A ensures that the position of each centromere is preserved even in post-mitotic cells as losing the centromere identity in such cells could be lethal in case of cell cycle re-entry since each chromosome would require a new centromere to form a functional kinetochore in that case of the need to resume cell division. Therefore the loss of centromere identity would force “de novo” centromere formation which could result in the activation/inactivation of genes such as tumor suppressors, transcription factors, oncogenes, cyclin dependent kinases etc. leading to cell death or in worst case scenario uncontrolled cell proliferation and cancer development.

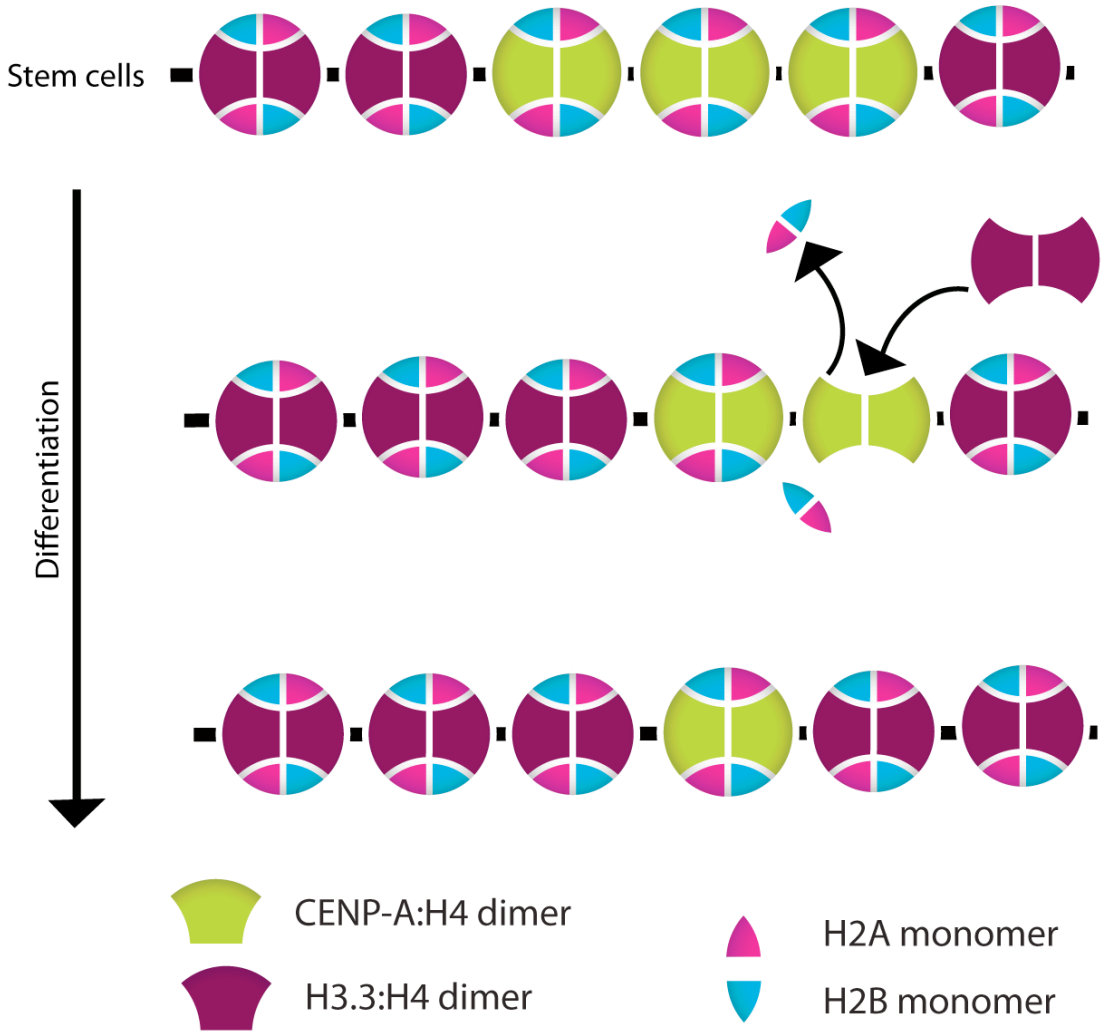

Figure 3: proposed model for nucleosome compositional changes at the centromeric chromatin upon terminal differentiation. In this model, a portion of CENP-A nucleosomes are disassembled and replaced with H3.3 nucleosomes in a step-wise manner to promote heterochromatinization in post-mitotic cells. H2A:H2B dimers are proposed to be released prior to the substitution of CENP-A:H4 dimers with H3.3:H4 dimers.

On the other hand, it would be of high interest to determine whether CENP-A is required for the process of differentiation. In this regard, shRNA depletion of CENP-A in differentiating C2C12 myoblasts could be used along with analysis of the expression of myogenic differentiation markers e.g. MyoD. If CENP-A is required for differentiation, this process would be compromised and most likely delayed due to the lack of CENP-A. Various parameters including cell viability, morphology, cell lineage markers and apoptosis could be assessed in these cells to characterize the role of CENP-A for the process of differentiation.

However, it is not clear whether other centromere proteins such as the constitutive centromere-associated network (CCAN) components CENP-C, CENP-T, CENP-W, CENP-S and CENP-X (Cheeseman and Desai, 2008; Hori et al., 2008; Hori et al., 2013) that are known to associate with the CENP-A nucleosome regardless of the cell cycle phase are retained at centromeres in differentiation. Therefore, it is highly warranted to investigate the colocalization of these proteins at post-mitotic centromeres. A variety of techniques including multi-color immunofluorescence could be used to this end.

References

Ambartsumyan, G., Gill, R.K., Perez, S.D., Conway, D., Vincent, J., Dalal, Y., and Clark, A.T. (2010). Centromere protein A dynamics in human pluripotent stem cell self-renewal, differentiation and DNA damage. Human molecular genetics *19*, 3970-3982.

Bui, M., Dimitriadis, E.K., Hoischen, C., An, E., Quenet, D., Giebe, S., Nita-Lazar, A., Diekmann, S., and Dalal, Y. (2012). Cell-Cycle-Dependent Structural Transitions in the Human CENP-A Nucleosome In Vivo. Cell *150*, 317-326.

Cheeseman, I.M., and Desai, A. (2008). Molecular architecture of the kinetochore-microtubule interface. Nature reviews Molecular cell biology *9*, 33-46.

Cleveland, D.W., Mao, Y., and Sullivan, K.F. (2003). Centromeres and kinetochores: from epigenetics to mitotic checkpoint signaling. Cell *112*, 407-421.

Codomo, C.A., Furuyama, T., and Henikoff, S. (2014). CENP-A octamers do not confer a reduction in nucleosome height by AFM. Nature structural & molecular biology *21*, 4-5.

Craig, J.M., Earnshaw, W.C., and Vagnarelli, P. (1999). Mammalian centromeres: DNA sequence, protein composition, and role in cell cycle progression. Experimental cell research *246*, 249-262.

Dalal, Y., Wang, H., Lindsay, S., and Henikoff, S. (2007). Tetrameric structure of centromeric nucleosomes in interphase Drosophila cells. PLoS biology *5*, e218.

Dernburg, A.F. (2001). Here, there, and everywhere: kinetochore function on holocentric chromosomes. The Journal of cell biology *153*, F33-38.

Dimitriadis, E.K., Weber, C., Gill, R.K., Diekmann, S., and Dalal, Y. (2010). Tetrameric organization of vertebrate centromeric nucleosomes. Proceedings of the National Academy of Sciences of the United States of America *107*, 20317-20322.

Dunleavy, E.M., Zhang, W., and Karpen, G.H. (2013). Solo or doppio: how many CENP-As make a centromeric nucleosome? Nature structural & molecular biology *20*, 648-650.

Furuyama, T., Codomo, C.A., and Henikoff, S. (2013). Reconstitution of hemisomes on budding yeast centromeric DNA. Nucleic acids research *41*, 5769-5783.

Gordon, M.S., Matei, D., Aghajanian, C., Matulonis, U.A., Brewer, M., Fleming, G.F., Hainsworth, J.D., Garcia, A.A., Pegram, M.D., Schilder, R.J.*, et al.* (2006). Clinical activity of pertuzumab (rhuMAb 2C4), a HER dimerization inhibitor, in advanced ovarian cancer: potential predictive relationship with tumor HER2 activation status. Journal of clinical oncology : official journal of the American Society of Clinical Oncology *24*, 4324-4332.

Hasson, D., Panchenko, T., Salimian, K.J., Salman, M.U., Sekulic, N., Alonso, A., Warburton, P.E., and Black, B.E. (2013). The octamer is the major form of CENP-A nucleosomes at human centromeres. Nature structural & molecular biology *20*, 687-695.

Henikoff, S., Ahmad, K., and Malik, H.S. (2001). The centromere paradox: stable inheritance with rapidly evolving DNA. Science *293*, 1098-1102.

Hori, T., Amano, M., Suzuki, A., Backer, C.B., Welburn, J.P., Dong, Y., McEwen, B.F., Shang, W.H., Suzuki, E., Okawa, K.*, et al.* (2008). CCAN makes multiple contacts with centromeric DNA to provide distinct pathways to the outer kinetochore. Cell *135*, 1039-1052.

Hori, T., Shang, W.H., Takeuchi, K., and Fukagawa, T. (2013). The CCAN recruits CENP-A to the centromere and forms the structural core for kinetochore assembly. The Journal of cell biology *200*, 45-60.

Howman, E.V., Fowler, K.J., Newson, A.J., Redward, S., MacDonald, A.C., Kalitsis, P., and Choo, K.H. (2000). Early disruption of centromeric chromatin organization in centromere protein A (Cenpa) null mice. Proceedings of the National Academy of Sciences of the United States of America *97*, 1148-1153.

Jain, A., Liu, R., Ramani, B., Arauz, E., Ishitsuka, Y., Ragunathan, K., Park, J., Chen, J., Xiang, Y.K., and Ha, T. (2011). Probing cellular protein complexes using single-molecule pull-down. Nature *473*, 484-488.

Jain, A., Liu, R., Xiang, Y.K., and Ha, T. (2012). Single-molecule pull-down for studying protein interactions. Nat Protoc *7*, 445-452.

Jin, C., and Felsenfeld, G. (2007). Nucleosome stability mediated by histone variants H3.3 and H2A.Z. Genes & development *21*, 1519-1529.

Karpen, G.H., and Allshire, R.C. (1997). The case for epigenetic effects on centromere identity and function. Trends in genetics : TIG *13*, 489-496.

Lavoie, H., Thevakumaran, N., Gavory, G., Li, J.J., Padeganeh, A., Guiral, S., Duchaine, J., Mao, D.Y., Bouvier, M., Sicheri, F.*, et al.* (2013). Inhibitors that stabilize a closed RAF kinase domain conformation induce dimerization. Nature chemical biology *9*, 428-436.

Maddox, P.S., Oegema, K., Desai, A., and Cheeseman, I.M. (2004). "Holo"er than thou: chromosome segregation and kinetochore function in C. elegans. Chromosome research

: an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology *12*, 641-653.

Miell, M.D., Fuller, C.J., Guse, A., Barysz, H.M., Downes, A., Owen-Hughes, T., Rappsilber, J., Straight, A.F., and Allshire, R.C. (2013). CENP-A confers a reduction in height on octameric nucleosomes. Nature structural & molecular biology *20*, 763-765.

Miell, M.D., Straight, A.F., and Allshire, R.C. (2014). Reply to "CENP-A octamers do not confer a reduction in nucleosome height by AFM". Nature structural & molecular biology *21*, 5-8.

Monen, J., Maddox, P.S., Hyndman, F., Oegema, K., and Desai, A. (2005). Differential role of CENP-A in the segregation of holocentric C. elegans chromosomes during meiosis and mitosis. Nature cell biology *7*, 1248-1255.

Padeganeh, A., Ryan, J., Boisvert, J., Ladouceur, A.M., Dorn, J.F., and Maddox, P.S. (2013). Octameric CENP-A nucleosomes are present at human centromeres throughout the cell cycle. Current biology : CB *23*, 764-769.

Pereira, S.L., Grayling, R.A., Lurz, R., and Reeve, J.N. (1997). Archaeal nucleosomes. Proceedings of the National Academy of Sciences of the United States of America *94*, 12633-12637.

Pereira, S.L., and Reeve, J.N. (1998). Histones and nucleosomes in Archaea and Eukarya: a comparative analysis. Extremophiles : life under extreme conditions *2*, 141-148.

Shivaraju, M., Unruh, J.R., Slaughter, B.D., Mattingly, M., Berman, J., and Gerton, J.L. (2012). Cell-cycle-coupled structural oscillation of centromeric nucleosomes in yeast. Cell *150*, 304-316.

Smith, M.M. (2002). Centromeres and variant histones: what, where, when and why? Current opinion in cell biology *14*, 279-285.

Sullivan, B.A., Blower, M.D., and Karpen, G.H. (2001). Determining centromere identity: cyclical stories and forking paths. Nature reviews Genetics *2*, 584-596.

Sullivan, L.L., Boivin, C.D., Mravinac, B., Song, I.Y., and Sullivan, B.A. (2011). Genomic size of CENP-A domain is proportional to total alpha satellite array size at human centromeres and expands in cancer cells. Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology *19*, 457-470.

Tachiwana, H., Kagawa, W., Shiga, T., Osakabe, A., Miya, Y., Saito, K., Hayashi-Takanaka, Y., Oda, T., Sato, M., Park, S.Y.*, et al.* (2011a). Crystal structure of the human centromeric nucleosome containing CENP-A. Nature *476*, 232-235.

Tachiwana, H., Osakabe, A., Shiga, T., Miya, Y., Kimura, H., Kagawa, W., and Kurumizaka, H. (2011b). Structures of human nucleosomes containing major histone H3 variants. Acta crystallographica Section D, Biological crystallography *67*, 578-583.

Torras-Llort, M., Moreno-Moreno, O., and Azorin, F. (2009). Focus on the centre: the role of chromatin on the regulation of centromere identity and function. The EMBO journal *28*, 2337-2348.

Verdaasdonk, J.S., and Bloom, K. (2011). Centromeres: unique chromatin structures that drive chromosome segregation. Nature reviews Molecular cell biology *12*, 320-332.

Walkiewicz, M.P., Dimitriadis, E.K., and Dalal, Y. (2014). CENP-A octamers do not confer a reduction in nucleosome height by AFM. Nature structural & molecular biology *21*, 2-3.

Yoda, K., Ando, S., Morishita, S., Houmura, K., Hashimoto, K., Takeyasu, K., and Okazaki, T. (2000). Human centromere protein A (CENP-A) can replace histone H3 in nucleosome reconstitution in vitro. Proceedings of the National Academy of Sciences of the United States of America *97*, 7266-7271.

Zhang, H., and Dawe, R.K. (2012). Total centromere size and genome size are strongly correlated in ten grass species. Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology *20*, 403-412.